Kongeriget Danmark

Patent application No.: 1156/96

Date of filing: 18 Oct 1996

Applicant: Dako A/S, Att: Marianne Hundewadt Christensen, Produktionsvej 42, 2600 Glostrup, DK

This is to certify the correctness of the following information:

The attached photocopy is a true copy of the following document:

- The specification, claims, abstract and drawings as filed with the application on the filing date indicated above.

Patentdirektoratet · Kongeriget Danmark ·

Erhvervsministeriet
Patentdirektoratet

TAASTRUP 09 Aug 1997

Karin Schlichting
Kontorfuldmægtig

NOVEL PROBES FOR THE DETECTION OF MYCOBACTERIA OF THE MYCOBACTERIUM TUBERCULOSIS COMPLEX

The present invention relates to the design, construction and use of novel probes for detecting mycobacteria of the Mycobacterium tuberculosis Complex (MTC) which probes are capable of detecting the organisms in test samples, e.g. expectorates, sputum, aspirates, urine, blood and tissue sections, food, soil and water.

BACKGROUND OF THE INVENTION

Tuberculosis caused by mycobacterial infection is presently the predominant infectious cause of morbidity and mortality world-wide, and is estimated to kill about three million people annually. WHO estimates that the annual number of new cases of tuberculosis will increase from 7.5 million in 1990 to 10.2 million in 2000, an escalation that will result in approximately 90 million new cases during this decade. It is furthermore estimated that 30 million people will die from tuberculosis during the 1990s, which equals one quarter of preventable deaths among adults.

The prevalence of tuberculosis has been very high in the poorer parts of the world such as Asia, Africa and South-America, but in recent years an increase has also been observed in industrialised countries. This appears to be due to an interaction of various factors including i.a. patterns of migration, poorly organised tuberculosis programmes and nutrition problems. Furthermore, a serious threat will arise from the emergence of new strains that are multi-drug resistant.

Considering the perspective and impact the disease has, the development of rapid, specific and preferably easy-performed and economic feasible diagnostic detection tests are of utmost importance and would be a very valuable tool in the fight against the spread of tuberculosis.

Generally, mycobacterial infections are divided into infections caused by two groups of bacteria, namely mycobacteria of the Mycobacterium tuberculosis Complex (MTC) and mycobacteria of the Mycobacterium avium-intracellulare Complex (MAC). The mycobacteria of the Mycobacterium tuberculosis Complex include M. tuberculosis, M. bovis and M. africanum, whereas the mycobacteria of the Mycobacterium avium-intracellulare Complex comprise M. avium and M. intracellulare.

Presently, the detection of mycobacteria by microscopy gives the more accurate diagnosis. The sample (e.g. an expectorate) is stained for the presence of acid-fast bacillus using Ziehl-

Neelsen staining and may subsequently be cultured in order to confirm the result obtained by staining. Such techniques are one of the cornerstones of all anti-tuberculosis programmes. However, the Ziehl-Neelsen staining lacks sensitivity since the detection limit is 10^4 organisms/ml or greater. On the contrary, cultivation is sensitive, and it may be possible to detect 10-100 organisms per sample, but the result is not available before up to 8 weeks of cultivation. Likewise, information of drug susceptibility is not available until after 2-3 weeks of further testing.

Automated detection is rapidly becoming available for large scale testing for the presence of mycobacteria. Such systems include ESP Myco Culture System (Difco), MB/BacT (Organon Teknika) and MGIT (Becton Dickinson). These test methods are based on colorimetric or fluorometric detection of carbon dioxide or oxygen produced by mycobacterial metabolism.

Neither staining nor cultivation methods allows distinction between the mycobacteria of the MTC and the MAC.

Some of the attempts to replace the methods based on cultivation rely on target amplification or target hybridisation using specific probes.

One of such newly developed target amplification method is based on PCR. The principle of this reaction is, through amplification of specific nucleic acid sequences of the mycobacteria, to increase the copy number of the specific sequence to a level where it may be detectable in an early stage of the infection. In principle, the PCR reaction offers the possibility of detecting as few as one target sequence. In most cases, the DNA is extracted prior to carrying out the PCR reaction. However, it has become clear that the method used to extract DNA from specimens has a great influence on the sensitivity and specificity of PCR products.

Furthermore, false negative results in specimens may be obtained due to the presence of inhibitors of the PCR reaction such as haemoglobin and proteins.

Another problem arises from cross-contamination of negative specimens with a bacteria not present in the sample. This may cause problems in conventional bacteriological procedures and may lead to a positive PCR result. Contamination of reagents and specimens with amplified PCR products is yet another well-recognised problem when using a PCR-based diagnosis.

Nucleic acid probes for detecting rRNA of mycobacteria have been described in for example US 5 547 842, EP-A 0 572 120 and US 5 422 242.

SUMMARY OF THE INVENTION

The present invention discloses and claims novel peptide nucleic acid probes for the detection of mycobacteria of the Mycobacterium tuberculosis Complex. The probes detect sequences in 23S rRNA and genomic sequences corresponding to said rRNA. rRNA is present in a high number of copies in each cell, and hence a well suited target for a sensitive test. Furthermore, probes that are complementary to said rRNA are especially suitable for hybridisation as it is known that species variable regions exist within these highly conserved sequences thereby enabling the design of probes for detecting mycobacteria of the Mycobacterium tuberculosis Complex.

The novel probes may be used in an assay for the detection of mycobacteria of the MTC group. The mycobacteria of the MTC group are responsible for significant morbidity and mortality in humans. M. tuberculosis is the most common mycobacteria of the MTC group isolated from humans. M. bovis may be transmitted from infected animals to humans. M. africanum causes pulmonary tuberculosis in tropical Africa.

Tuberculosis is highly contagious, and a rapid diagnosis of the disease is therefore very important. For most clinical laboratories, assignment of an isolate to the group of MTC bacteria is sufficient.

Thus, in a first aspect, the invention features a hybridisation assay probe able to detect mycobacteria of the MTC group. Specifically, the probe is a peptide nucleic acid as defined in claim 1. Such probes do not to any significant degree cross react with nucleic acid from other organisms in the test sample under appropriate stringency conditions.

In another aspect, the present invention relates to a method according to claim 7 for detecting the presence of organisms belonging to the MTC group.

In yet another aspect, the present invention relates to a kit comprising at least one peptide nucleic acid probe as defined in anyone of claims 1 to 6.

BRIEF DESCRIPTION OF THE FIGURES

Figure 1 shows positions 281 to 400 and 601 to 680 of an alignment of 23S rRNA of M. tuberculosis (positions 1251 to 1368 and 1568 to 1647 of GenBank entry GB:MTCY130, accession number Z73992), M. avium (positions 579 to 697 and 896 to 960 of GenBank entry

GB:MA23SRNA, accession number X74494), M. phlei (positions 665 to 783 and 978 to 1047 of GenBank entry GB:MP23SRNA, accession number X74493), M. leprae (positions 591 to 709 and 909 to 973 of GenBank entry GB:ML5S23S, accession number X56657), M. paratuberculosis (positions 579 to 697 and 896 to 960 of GenBank entry GB:MPARRNA, accession number X74495), M. gastri (positions 366 to 484 and 684 to 748 of GenBank entry GB:MG23SRRNA, accession number Z17211) and M. kansasii (positions 309 to 427 and 627 to 691 of GenBank entry GB:MK23SRRNA, accession number Z17212). Preferred nucleobase sequences of probes may be chosen within positions 326 to 359 or 635 to 658 of the alignment corresponding to positions 1296 to 1327 and 1602 to 1625 of M. tuberculosis sequence (indicated by heavy frames). Mismatches between the sequences of M. avium, M. phlei, M. leprae, M. paratuberculosis, M. gastri and M. kansasii and that of M. tuberculosis in the alignment are indicated by light frames.

Figure 2 shows positions 721 to 840 of an alignment of 23S rRNA of M. tuberculosis (positions 1688 to 1807 of GenBank entry GB:MTCY130, accession number Z73992), M. avium (positions 1001 to 1109 of GenBank entry GB:MA23SRNA, accession number X74494), M. phlei (positions 1088 to 1203 of GenBank entry GB:MP23SRNA, accession number X74493), M. leprae (positions 1014 to 1125 of GenBank entry GB:ML5S23S, accession number X56657), M. paratuberculosis (positions 1001 to 1109 of GenBank entry GB:MPARRNA, accession number X74495), M. gastri (positions 789 to 900 of GenBank entry GB:MG23SRRNA, accession number Z17211) and M. kansasii (positions 732 to 843 of GenBank entry GB:MK23SRRNA, accession number Z17212). Preferred nucleobase sequences of probes may be chosen within positions 761 to 787 of the alignment corresponding to positions 1728 to 1754 of M. tuberculosis sequence (indicated by a heavy frame). Mismatches between the sequences of M. avium, M. phlei, M. leprae, M. paratuberculosis, M. gastri and M. kansasii and that of M. tuberculosis in the alignment are indicated by light frames.

Figure 3 shows positions 1281 to 1360 and 1601 to 1680 of an alignment of 23S rRNA of M. tuberculosis (positions 2246 to 2323 and 2563 to 2636 of GenBank entry GB:MTCY130, accession number Z73992), M. avium (positions 1549 to 1626 and 1865 to 1938 of GenBank entry GB:MA23SRNA, accession number X74494), M. phlei (positions 1643 to 1720 and 1960 to 2027 of GenBank entry GB:MP23SRNA, accession number X74493), M. leprae (positions 1565 to 1644 and 1884 to 1959 of GenBank entry GB:ML5S23S, accession number X56657), M. paratuberculosis (positions 1549 to 1626 and 1865 to 1938 of GenBank entry GB:MPARRNA, accession number X74495), M. gastri (positions 1339 to 1406 and 1646 to 1719 of GenBank GB:MG23SRRNA, accession number Z17211) and M. kansasii (positions 1282 to 1349 and 1589 to 1663 of GenBank entry GB:MK23SRRNA, accession number

Z17212). Preferred nucleobase sequences of probes may be chosen within positions 1306 to 1322 and 1621 to 1631 of the alignment corresponding to positions 2271 to 2285 and 2581 to 2591 of M. tuberculosis sequence (indicated by heavy frames). Mismatches between the sequences of M. avium, M. phlei, M. leprae, M. paratuberculosis, M. gastri and M. kansasii and that of M. tuberculosis in the alignment are indicated by light frames.

Figure 4 shows positions 2361 to 2520 and 3081 to 3120 of an alignment of 23S rRNA of M. tuberculosis (positions 3307 to 3466 and 4026 to 4064 of GenBank entry GB:MTCY130, accession number Z73992), M. avium (positions 2607 to 2765 and 3325 to 3363 of GenBank entry GB:MA23SRNA, accession number X74494), M. phlei (positions 2699 to 2858 ans 3418 to to 3456 of GenBank entry GB:MP23SRNA, accession number X74493), M. leprae (positions 2630 to 2789 of GenBank entry GB:ML5S23S, accession number X56657), M. paratuberculosis (positions 2607 to 2765 and 3325 to 3363 of GenBank entry GB:MPARRNA, accession number X74495), and M. kansasii (positions 2334 to 2494 and 3053 to 3091 of GenBank entry GB:MK23SRRNA, accession number Z17212). Preferred nucleobase sequences of probes may be chosen within positions 2401 to 2418, 2455 to 2486 and 3094 to 3103 of the alignment corresponding to positions 3347 to 3364, 3401 to 3432 and 4038 to 4047 of M. tuberculosis sequence (indicated by heavy frames). Mismatches between the sequences of M. avium, M. phlei, M. leprae, M. paratuberculosis, M. gastri and M. kansasii and that of M. tuberculosis in the alignment are indicated by light frames.

SPECIFIC DESCRIPTION

The present invention provides novel probes for use in rapid and sensitive hybridisation based assays for the detection of organisms belonging to the MTC group.

We have identified suitable variable regions of the target nucleic acid by comparative analysis of generally available 23S rRNA sequences. Computers and computer programs which have been used for the purposes herein disclosed are generally available. In the probe design, sequence variations between the organisms belonging to the MTC group and other organisms have been taken into consideration, in particular M. avium.

When designing the probes, due regard should be taken to the assay conditions under which the probes are to be used. The stringency of the assay conditions determines the degree of complementarity needed between the probe and nucleic acid forming a hybrid. Stringency is chosen so as to maximise the difference in stability between the hybrid formed with the target nucleic acid and the non-target nucleic acid. It is desirable to have probes which hybridise under conditions of high stringency. Under such conditions, only highly complementary nucleic

acids will form stable hybrids with the probe according to the invention; stable hybrids without a sufficient degree of complementarity will not be formed.

Furthermore, probes should be positioned so as to minimise the stability of the probe:non-target nucleic acid hybrid. This may be accomplished by minimising the degree of complementarity to non-target nucleic acid and by designing the probe to span as many destabilising mismatches as possible. Whether a probe is useful to detect an organism belonging to the MTC group depends largely on the thermal stability difference between probe:target hybrids and probe:non-target hybrids. In designing the probes, the differences in these T_m values should be as large as possible.

Hybrids formed between peptide nucleic acid probes and nucleic acids have a higher thermal instability of mismatching bases compared to nucleic acid duplexes of the same sequences. Thus, the peptide nucleic acid probes exhibit a greater specificity for a complementary nucleic acid sequence than the traditional nucleic acid probe, which is seen as a greater difference in T_m values for probe:target hybrids and probe:non-target hybrids.

The length of the probe sequence is also important. The optimal length of a probe comprising a particular site of differences in base composition, e.g. among homologous regions of mycobacteria 23S rRNA, is a compromise between the principle that longer probes ensure specificity and shorter probes ensure that the destabilising differences in base composition constitute a greater portion of the probe.

Peptide nucleic acids can form duplexes in either orientation, but the antiparallel orientation form the most regular and stable duplex. Hence the antiparallel configuration is preferred for probe applications.

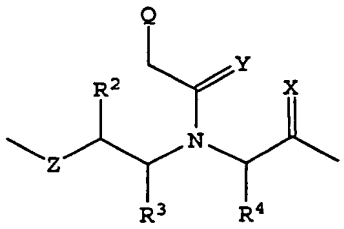
Mainly because the peptide nucleic acid strand is uncharged, a peptide nucleic acid-nucleic acid-duplex will have a higher T_m than the corresponding nucleic acid-nucleic acid-duplex. Typically there will be an increase in T_m of about 1 °C per base pair at 100 mM NaCl depending on the sequence (Egholm et al. (1993), Nature, 365, 566-568).

In contrast to DNA-DNA-duplex formation, no salt is necessary to facilitate and stabilise the formation of a peptide nucleic acid-DNA or a peptide nucleic acid-RNA duplex. The T_m of the peptide nucleic acid-DNA-duplex changes only a little with increasing ionic strength. Typically for a 15-mer, the T_m will drop only 5 °C when the salt concentration is raised from 10 mM NaCl to 1 M NaCl. At low ionic strength (e.g. 10 mM phosphate buffer with no salt added), it is possible to hybridise peptide nucleic acid to a target sequence under conditions where no

stable DNA-DNA-duplex formation is able to occur (Nucleic Acid Hybridisation, a practical approach, eds. B. D. Hames & S. J. Higgins, IRL Press 1985, page 62-64). Furthermore, target sites that normally are inaccessible can be made more readily accessible for hybridisation with peptide nucleic acid probes at low salt concentration as the secondary and tertiary structure of nucleic acids are melted under such conditions.

Although it is preferred to use peptide nucleic acid probes targeting specific sequences of rRNA, it will readily be understood that peptide nucleic acid probes complementary to the rRNA targeting probes will be useful for the detection of the genes (DNA) coding for said sequence specific rRNA. Thus, as used herein, "probes able to form hybrids with target sequences in 23S rRNA" refers to probes capable of hybridising to sequences in 23S rRNA or to corresponding sequences in the non-coding strand of the rDNA as well as it refers to complementary probes capable of hybridising to the coding strand of DNA coding for the target rRNA sequences.

In accordance with the present invention, peptide nucleic acid probes of formula (I) are provided, which probes are useful for detecting mycobacteria of the Mycobacterium tuberculosis Complex (MTC) in a sample, and which probes comprise from 10 to 30 polymerised moieties of formula (I)



(I)

wherein each X and Y independently designate O or S,
each Z independently designates O, S, NR^1, or $C(R^1)_2$, wherein each R^1 independently designate H, C_{1-6} alkyl, C_{1-6} alkenyl, C_{1-6} alkynyl,
each R^2, R^3 and R^4 independently designate H, the side chain of a naturally occurring amino acid, the side chain of a non-naturally occurring nucleobase, C_{1-4} alkyl, C_{1-4} alkenyl or C_{1-4} alkynyl, or a functional group, each Q independently designates a naturally occurring nucleobase, a non-naturally occurring nucleobase, an intercalator, a nucleobase-binding group, a label or H,

with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase complementary to a nucleobase of M. tuberculosis 23S rRNA that differs from the corresponding nucleobase of M. avium located

within the following domains

Position 326 to Position 359 in Figure 1, or
Position 635 to Position 658 in Figure 1, or
Position 761 to Position 787 in Figure 2, or
Position 1306 to Position 1322 in Figure 3, or
Position 1621 to Position 1631 in Figure 3, or
Position 2401 to Position 2418 in Figure 4, or
Position 2455 to Position 2486 in Figure 4, or
Position 3094 to Position 3103 in Figure 4,

and further with the proviso that the probe comprising such subsequence is able to form hybrids with target sequences in 23S rRNA of said mycobacteria.

The term "naturally occurring nucleobases" includes the four main DNA bases (i.e. thymine (T), cytosine (C), adenine (A) and guanine (G)) as well as other naturally occurring nucleobases (e.g. uracil (U) and hypoxanthine).

The term "non-naturally occurring nucleobases" comprises i.a. modified naturally occurring nucleobases. Such non-naturally occurring nucleobases may be modified by substitution by e.g. one or more C_{1-8} alkyl, C_{1-8} alkenyl or C_{1-8} alkynyl groups or labels. Examples of non-naturally occurring nucleobases are purine, 2,6-diamino purine, 5-propynylcytosine (C propynyl), isocytosine (iso-C), 5-methyl-isocytosine ($iso^{Me}C$) (see e.g. Tetrahedron Letters Vol 36, No 12, 2033-2036 (1995) or Tetrahedron Letters Vol 36, No 21, 3601-3604 (1995)), 7-deazaadenine, 7-deazaguanine, N^4-ethanocytosine, N^6-ethano-2,6-diaminopurine, 5-(C_{3-6})-alkenyluracil, 5-(C_{3-6})-alkynylcytosine, 5-fluorouracil and pseudocytosine.

Examples of useful intercalators are e.g. acridin, antraquinone, psoralen and pyrene.

Examples of useful nucleobase-binding groups are e.g. groups containing cyclic or heterocyclic rings. Non-limiting examples are 3-nitro pyrrole and 5-nitro indole.

It is to be understood that alkyl, alkenyl and alkynyl groups may be branched or non-branched, cyclic or non-cyclic, and may further be interrupted by one or more heteroatoms, or may be unsubtituted or substituted by or may contain one or more functional groups. Non-limiting examples of such functional groups are acetyl groups, acyl groups, amino groups, carbamido groups, carbamoyl groups, carbamyl groups, carbonyl groups, carboxy groups, cyano groups, dithio groups, formyl groups, guanidino groups, halogens, hydrazino groups, hydrazo groups,

hydroxamino groups, hydroxy groups, keto groups, mercapto groups, nitro groups, phospho groups, phospho ester groups, sulfo groups, thiocyanato groups, cyclic, aromatic and heterocyclic groups.

C_{1-4} groups contain from 1 to 4 carbon atoms, C_{1-6} groups contain from 1 to 6 carbon atoms, and C_{1-15} groups contain from 1 to 15 carbon atoms, not including optional substituents, heteroatoms and/or functional groups. Non-limiting examples of such groups are -OH, $-CH_3$, $-CF_3$, $-CH_2-$, $-CH_2CH_3$, $-CH_2CH_2-$, $-CH(CH_3)_2$, $-OCH_3$, $-OCH_2-$, $-OCH_2CH_3$, $-OCH_2CH_2-$, $-OCH(CH_3)_2$, $-OC(O)CH_3$, $-OC(O)CH_2-$, -C(O)H, -C(O)-, $-C(O)CH_3$, -C(O)OH, -C(O)O-, $-CH_2NH_2$, $-CH_2NH-$, $-CH_2OCH_3$, $-CH_2OCH_2-$, $-CH_2OC(O)OH$, $-CH_2OC(O)O-$, $-CH_2C(O)CH_3$, $-CH_2C(O)CH_2-$, $-C(O)NH_2$, $-P(O)_4H$, -SH, $-NH_2$, $-CH=CH_2$, -CH=CH-, $-CH=CHCH_2C(O)OH$, $-CH=CHCH_2C(O)O-$, $-C\equiv CH$, $-C\equiv C-$, $-CH_2C\equiv CH$, $-CH_2C\equiv C-$, $-CH_2C\equiv CCH_3$, $-OCH_2C\equiv CH$, $-OCH_2C\equiv C-$, $-OCH_2C\equiv CCH_3$, $-NHCH_2C(O)-$, $-NHCH_2CH_2C(O)-$, $-NH(CH_2CH_2O)_2CH_2C(O)-$, and $HO(O)CCH_2C(O)(NH-(CH_2CH_2O)_2CH_2C(O))_2-$, phenyl, benzyl, naphthyl, oxazolyl, pyridinyl, thiadiazolyl, triazolyl, and thienyl.

Within the present context, the expression "naturally occurring amino acid" is intended to comprise D- and L-forms of amino acids commonly found in nature, e.g. D- and L-forms of Ala (alanine), Arg (arginine), Asn (aspargine), Asp (aspartic acid), Cys (cysteine), Gln (glutamine), Glu (glutamic acid), His (histidine), Ile (isoleucine), Leu (leucine), Lys (lysine), Met (methionine), Phe (phenylalanine), Pro (proline), Ser (serine), Thr (threonine), Trp (tryptophan), Tyr (tyrosine) and Val (valine).

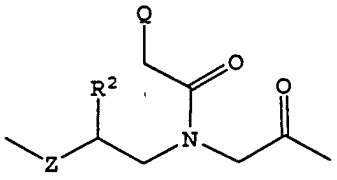
In the present context, the expression "non-naturally occurring amino acid" is intended to comprise D- and L-forms of amino acids other than those commonly found in nature as well as modified naturally occurring amino acids. Examples of useful non-naturally occurring amino acids are D- and L-forms of Cha (cyclohexylalanine), Cit (citrulline), Hci (homocitrulline), HomoCys (homocystein), Hse (homoserine), Nle (norleucine), Nva (norvaline), Orn (ornithine), Sar (sarcosine) and Thi (thienylalanine).

The strength of the binding between the probe and the nucleic acid sequence is further influenced by the ligand Q. When Q designates a nucleobase, Hoogsteen and/or Watson-Crick base pairing assists in the formation of hybrids between a nucleic acid sequence to be detected and the probe. It is contemplated that one or more of the ligands may be a group which contribute little or none to the binding of the nucleic acid such as hydrogen. It is contemplated that suitable probes to be used comprise less than 25% by weight of moieties, wherein Q designates such groups. One or more of the ligands Q may be groups that stabilise nucleobase stacking such as intercalators or nucleobase-binding groups.

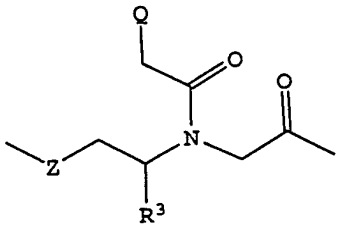
In the above-indicated probes one or more of the Q-groups may designate a label. Examples of suitable labels are given below. Moieties wherein Q denotes a label may preferably be located in one or both of the terminating moieties of the probe. Moieties wherein Q denotes a label may also be located internally.

The peptide nucleic acid probes may comprise moieties, wherein all X groups are O (polyamides) or wherein all X groups are S (polythioamides). It is to be understood that the probes may also comprise mixed moieties (comprising both amide and thioamide moieties).

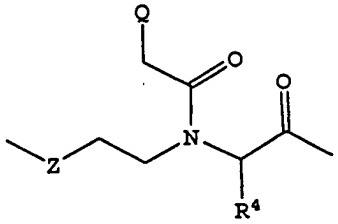
In another aspect, the present invention relates to peptide nucleic acid probes of formula (II), (III) and (IV)



(II)



(III)



(IV)

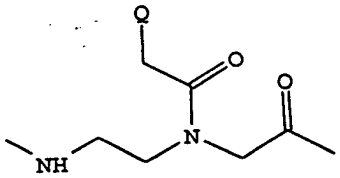
wherein Z, R^2, R^3, and R^4, and Q is as defined above, which probes are suitable for detecting mycobacteria of the MTC.

In a preferred embodiment, the peptide nucleic acid probes according to the invention are of formulas (I)-(IV) as defined above with Z being NH, NCH_3 or O, each R^2, R^3 and R^4 independently being the side chain of a naturally occurring nucleobase, the side chain of a non-naturally occurring nucleobase, or C_{1-4} alkyl, and each Q being a naturally occurring nucleobase or a non-naturally occurring nucleobase.

Peptide nucleic acid probes according to the invention are preferably those of formula (I)-(IV)

as defined above, wherein Z is NH or O, and R^2 is H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, and Q is a nucleobase selected from thymine, adenine, cytosine, guanine, uracil, iso-C, iso-G and 2,6-diaminopurine.

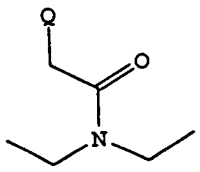
Peptide nucleic acid probes, which are of major interest for detecting mycobacteria of the MTC group, are probes of formula (V)



(V)

wherein R^4 is H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, and Q is as defined above and with the provisos indicated above.

The peptide nucleic acid probe comprises polymerised moieties as defined above. From the formula, it is to be understood that the probe may comprise polymerised moieties which structure may be mutually different or identical. It may be advantageous that at least one moiety of the probe, preferably one (or both) of the moieties terminating the probe, are of a different structure. Such terminating moieties may suitably be a moiety of formula (VI)



(VI)

where Q is as defined above. Such moiety may suitably be connected to a peptide nucleic acid moiety though an amide bond.

The preferred length of the probe will depend on the target material and whether labelled probes are used. It is contemplated that especially interesting probes comprise from 10 to 30 polymerised moieties as defined above. Probes of the invention may suitably comprise from 12 to 25 polymerised moieties, more suitably from 14 to 22 polymerised moieties, most suitably from 15 to 20 polymerised moieties.

As mentioned above, the polymerised moieties of the probes may be mutually different or identical. In some embodiments, the polymerised moieties of formulas (V) constitute at least 75% by weight (calculated by excluding labels and linkers), preferably at least 80% by weight and most preferably at least 90% by weight of the probe.

The ends on the moieties terminating the probe may be substituted by suitable substituents which in the following will be named "linkers". A terminating end may suitably be substituted by from 1 to 5 linkers, more suitably from 1 to 3 linkers. Such linkers may suitably be selected among C_{1-15} alkyl, C_{1-15} alkenyl and C_{1-15} alkynyl groups as defined above. The linkers may be substituted or unsubstituted, branched or non-branched, or be interrupted by heteroatoms, or be substituted or contain functional groups as described above. This may depend on the chemical nature of the terminating moiety (i.e. whether the moiety is terminated by a carbon, oxygen or nitrogen atom). A terminating end or a linker on a terminating end may further be substituted by one or more labels, which labels may be incorporated end to end, i.e. so as to form a non-branched labelled end, or may be incorporated so as to form a branched labelled end ("zipper"). The linkers may be attached directly to a terminating end, may be attached to a label or between labels on a terminating end, or be attached to a terminating end before a label is attached to a terminating end. It should be understood that two terminating ends may carry different or identical substituents, linkers and/or labels. It should further be understood that the term "a label" is intended to comprise one or more labels as the term "linkers" is to comprise one or more linkers.

Examples of suitable linkers are $-NH(CH_2CH_2O)_nCH_2C(O)-$, $-NH(CHOH)_nC(O)-$, $-(O)C(CH_2OCH_2)_nC(O)-$ and $-NH(CH_2)_nC(O)-$, $NH_2(CH_2CH_2O)_nCH_2C(O)-$, $NH_2(CHOH)_nC(O)-$, $HO(O)C(CH_2OCH_2)_nC(O)-$, $NH_2(CH_2)_nC(O)-$, $-NH(CH_2CH_2O)_nCH_2C(O)OH$, $-NH(CHOH)_nC(O)OH$, $-(O)C(CH_2OCH_2)_nC(O)OH$ and $-NH(CH_2)_nC(O)OH$, wherein n is 0 or an integer from 1 to 8, preferably from 1 to 3. Examples of very interesting linkers are $-NHCH_2C(O)-$, $-NHCH_2CH_2C(O)-$, $-NH(CH_2CH_2O)_2CH_2C(O)-$, $HO(O)CCH_2CH_2C(O)(NH-(CH_2CH_2O)_2CH_2C(O))_2-$.

In the present context, the term "label" refers to a substituent which is useful for detection or visualisation. Suitable labels comprise fluorophores, biotin, dinitro benzoic acid, digoxigenin, radioisotope labels, peptide or enzyme labels, chemiluminiscence labels, hapten, antigen or antibody labels.

The expression "peptide label" is intended to mean a label comprising from 1 to 20 naturally occurring or non-naturally occurring amino acids, preferably from 1 to 10 naturally occurring or non-naturally occurring amino acids, more preferably from 1 to 8 naturally occurring or non-naturally occurring amino acids, most preferably from 1 to 4 naturally occurring or non-naturally occurring amino acids, linked together end to end in a non-branched or branched ("zipper") fashion. In a preferred embodiment, such a non-branched or branched end comprises one or more, preferably from 1 to 8 labels, more preferably from 1 to 4, further labels other than a peptide label. Such further labels may suitably terminate a non-branched

end or a branched end. One or more linkers may suitably be attached to the terminating end before a peptide label and/or a further label is attached. Such linker units may also be attached between a peptide label and a further label.

The probe as such may also comprise one or more labels such as from 1 to 8, preferably from 1 to 4, labels and/or one or more linker units, which may be attached internally, i.e. to the backbone of the probe. The linker units and labels may mutually be attached as described above.

Examples of particular interesting labels are biotin, fluorescent labels, such as fluorescein labels, e.g. 5-(and 6)-carboxyfluorescein, 5- or 6-carboxyfluorescein, 6-(fluorescein)-5-(and 6)-carboxamido hexanoic acid and fluorescein isothiocyanate, peptide labels consisting of from 1 to 20 naturally occurring amino acids or non-naturally occurring amino acids, peroxidases such as horse radish peroxidase (HRP) and soya bean peroxidase, dinitro benzoic acid, rhodamine, tetramethylrhodamine, cyanine dyes such as Cy2, Cy3 and Cy5, coumarin, R-phycoerythrin (RPE), allophycoerythrin, Texas Red and Princeton Red as well as conjugates of R-phycoerythrin and, e.g. Cy5 or Texas Red.

Examples of preferred labels are biotin, fluorescent labels, peptide labels and dinitro benzoic acid. Peptide labels may preferably be composed of from 1 to 10, more preferably of from 1 to 8, most preferably of from 1 to 4, naturally occurring or non-naturally occurring amino acids. It may be particularly advantageous to incorporate one or more labels other as well as a peptide label such as from 1 to 8 or from 1 to 4 other labels.

Suitable peptide labels may preferably be composed of cysteine, glycine, lysine or ornithine.

In a further embodiment, the Q substituent as defined above may be labelled. Suitable labels are as defined above. Between Q and such a label, a linker as defined above may be incorporated. It is preferred that such labelled ligands Q are selected from thymine and uridine labelled in the 5-position and 7-deazaguanine and 7-deazaadenine labelled in the 7-position.

The probes may be synthesised according to the procedures described in "PNA Information Package" obtained from Millipore Corporation (Bedford, MA, USA), or may be synthesised on an Expedite Nucleic Acid Synthesis System (PerSeptive, USA).

If using the Fmoc strategy for elongation of the probe with linkers or amino acids, it was possible to retain side chain amino groups protected with acid sensitive protection groups such as the Boc or Mtt group. This method allows introduction of a linker containing several

Boc protected amino groups which can all be cleaved and labelled in the same synthesis cycle.

One way of labelling a probe is to use a fluorescent label, such as 5-(and 6)-carboxyfluorescein, 5- or 6-carboxyfluorescein, or 6-(fluorescein)-5-(and 6)-carboxamido hexanoic acid. The acid group is activated with HATU (O-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) or HBTU (2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) and reacted with the N-terminal amino group. The same technique can be applied to other labelling groups containing an acid function. Alternatively, the succinimidyl ester of the above-mentioned labels or fluorescein isothiocyanate may be used directly.

After synthesis, probes were cleaved from the resin using standard procedures as described by Millipore Corporation or PerSeptive Biosystems. The probes were purified and analysed using reversed-phase HPLC techniques at 50°C and were characterised by matrix-assisted laser desorption/ionisation time of flight mass spectrometry (MALDI-TOFMS), plasma desorption mass spectrometry (PDMS), electron spray mass spectrometry (ESMS), or fast atom bombardment (FAB-MS).

Generally, probes such as probes comprising polymerised moieties of formula (V) may also be prepared as described in, e.g., Tetrahedron Letters Vol 35, No 29, 5173-5176 (1994) and Bioorganic & Medical Chemistry Letters, Vol 4, No 8, 1077-1080 (1994). Chemical properties of some probes are described in, e.g., Nature, 365, 566-568 (1993).

Detection of the label depend on the type of label and on the format of the procedure. In cases where the sample is deposited onto slides, the hybridisation results may be visualised using well known immunohistochemical staining methods to detect the labelling on the probe. When fluorescent labelled binding partners are used, the hybrids may be detected using an antibody against the fluorescent label which antibody may be conjugated with an enzyme. The fluorescent label may alternatively be detected directly using a fluorescence microscope, or the results may be automatically analysed on a fluorescent-based image analysis system.

When biotin labelled probes are used, the hybrids may be detected using an antibody against the biotin label which antibody may be conjugated with an enzyme. If necessary, an enhancement of the signal can be generated using commercially available amplification systems such as the catalysed signal amplification system for biotinylated probes (DAKO K 1500).

The probes according to the invention are used in the detection of mycobacteria of the MTC in

samples which may contain these bacteria.

In the assay method, at least one probe according to the invention is contacted with target nucleic acid and an analysis for hybrid formation is carried out.

In the assay method according to the invention, a sample to be analysed for the presence of mycobacteria of the MTC is contacted with one or more probes according to the invention under such conditions by which hybridisation between the probe and any complementary sample rRNA of mycobacteria of the MTC can occur, and observing or measuring the resulting hybridisation.

In one embodiment of the assay method, conventionally prepared smears of bacterial cells are contacted with one or more probes according to the invention under conditions suitable for hybridisation to occur between the probe(s) and any complementary rRNA in the sample. The complexes formed are detected. An example of this assay format is fluorescence *in situ* hybridisation (FISH), wherein the probes according to the invention are labelled with fluorescein or another fluorophore. When designing MTC probes, it might be advantageous to use more than one probe. If e.g. three such probes are included in the assay each in a concentration of one third of the concentration of a single probe, possible cross reactivity of the individual probes will not invalidate the results.

In another embodiment of the assay method, a test sample is firstly subjected to conditions, which release nucleic acid from the bacteria present in that sample. Contact between one or more probes as defined herein, which may be labelled, and the rRNA target may be carried out in solution under conditions, which promote hybridisation between the probe(s) and any target nucleic acid present. The probe:nucleic acid complex may be immobilised to a solid support, e.g. by using a capture probe.

Due to the high affinity of the probes defined herein for nucleic acids, it is not necessary to carry out the hybridisation of the probe and nucleic acid in solution. This allows flexibility in the assay format. For instance, the detection probes can be brought into contact with the target nucleic acid in solution and the probe/nucleic acid complex can be captured by an immobilised capture probe. Or the sample comprising the target nucleic acid can even be added to an assay system comprising detection probes as well as immobilised capture probe. The immobilisation of the capture probe may be effected by using a streptavidin coated solid phase and a biotinylated capture probe. The probe may be immobilised onto a solid support by coupling reaction between a carboxylic acid on the linker and an amino derivatised support. Alternatively, the coupling onto the solid support may be accomplished by photochemical

activation of photoreactive groups which have been attached absorptively to the solid support prior to photochemical activation. Such photoreactive groups are described in EP 408 078 A.

In practice, a solid phase based assay system is very attractive as the analysis can be carried out using a solid phase precoated with a capture probe. A solid phase based assay system is also feasible for automatisation of the analysis.

The capture probe may be one of the other MTC probes not used in the hybridisation reaction and detection step for target nucleic acid, thus ensuring dual species specificity. The dual specificity will allow shorter probes be used, e.g. 10 mer probes.

The solid support capture system may take a wide variety of forms well known in the art, such as e.g. a plate, a microtiter plate having one or more wells, a microscope slide, a filter, a membrane, a tube, a dip stick, a strip, beads such as paramagnetic beads, beads made of polystyrene, polypropylene, polyethylene, dextran, nylon, amyloses, natural and modified celluloses, polyacrylamides and agaroses. When a filter, a membrane, a strip or beads is (are) used as the solid support, it (they) may, if conveniently be incorporated into a single-use.

It has been observed that peptide nucleic acids may bind to a variety of solid phases. A blocking reaction is required to reduce non-specific binding of the peptide nucleic acids to the solid phase. The blocking reaction may be carried out with commonly used blocking reagents, such as BSA (bovine serum albumin), casein, Triton X-100® or Tween 20®. The preferred blocking reagents are Triton X-100® and Tween 20®.

The captured probe:nucleic acid complexes may be detected or identified by a wide variety of methods for that purpose. The probe to be brought in contact with the target nucleic acid may be labelled, whereby said may form part of the detection system. In another embodiment, the captures probe:nucleic acid complexes are detected using a detection system based on an antibody reacting specifically with complexes formed between peptide nucleic acid and nucleic acid (such as described in WO 95/17430), in which detection system the primary antibody may comprise a label, or which detection system comprises a labelled secondary antibody, which specifically binds to the primary antibody.

DESCRIPTION OF SPECIFIC EMBODIMENTS

Examples of suitable Qs of adjacent moieties are given below. Peptide nucleic acid probes comprising such Qs will be able to detect mycobacteria of the MTC group.The probes are written from left to right corresponding to from the C-terminal end towards the N-terminal end.

The nucleobase mismatche(s) between M. tuberculosis and M. avium is (are) indicated in bold.

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 326 to 359 shown in Figure 1 are the following

AG**C** GCT GAG ACA TAT **CC**T CCC A
G**C** GCT GAG ACA TAT **CC**T CCC A
C GCT GAG ACA TAT **CC**T CCC A
AG**C** GCT GAG ACA TAT **CC**T CCC
AG**C** GCT GAG ACA TAT **CC**T CC
AG**C** GCT GAG ACA TAT **CC**T C
AG**C** GCT GAG ACA TAT **CC**T
AG**C** GCT GAG ACA TAT **CC**

TAG **C**GC TGA GAC ATA T**CC** TCC C
AG **C**GC TGA GAC ATA T**CC** TCC C
G **C**GC TGA GAC ATA T**CC** TCC C
CGC TGA GAC ATA T**CC** TCC C
TAG **C**GC TGA GAC ATA T**CC** TCC
TAG **C**GC TGA GAC ATA T**CC** TC
TAG **C**GC TGA GAC ATA T**CC** T
TAG **C**GC TGA GAC ATA T**CC**

GTA G**C**G CTG AGA CAT AT**C C**TC C
TA G**C**G CTG AGA CAT AT**C C**TC C
A G**C**G CTG AGA CAT AT**C C**TC C
G**C**G CTG AGA CAT AT**C C**TC C
CG CTG AGA CAT AT**C C**TC C
GTA G**C**G CTG AGA CAT AT**C C**TC
GTA G**C**G CTG AGA CAT AT**C C**T
GTA G**C**G CTG AGA CAT AT**C C**

GGT AG**C** GCT GAG ACA TAT **CC**T C
GT AG**C** GCT GAG ACA TAT **CC**T C
T AG**C** GCT GAG ACA TAT **CC**T C
AG**C** GCT GAG ACA TAT **CC**T C
G**C** GCT GAG ACA TAT **CC**T C

C GCT GAG ACA TAT **CC**T C
GGT AG**C** GCT GAG ACA TAT **CC**T
GGT AG**C** GCT GAG ACA TAT **CC**

GGG TAG **C**GC TGA GAC ATA T**CC** T
GG TAG **C**GC TGA GAC ATA T**CC** T
G TAG **C**GC TGA GAC ATA T**CC** T
TAG **C**GC TGA GAC ATA T**CC** T
AG **C**GC TGA GAC ATA T**CC** T
G **C**GC TGA GAC ATA T**CC** T
CGC TGA GAC ATA T**CC** T
GGG TAG **C**GC TGA GAC ATA T**CC**

C**G**G GTA G**C**G CTG AGA CAT AT**C C**
GG GTA G**C**G CTG AGA CAT AT**C C**
G GTA G**C**G CTG AGA CAT AT**C C**
GTA G**C**G CTG AGA CAT AT**C C**
TA G**C**G CTG AGA CAT AT**C C**
A G**C**G CTG AGA CAT AT**C C**
G**C**G CTG AGA CAT AT**C C**
CG CTG AGA CAT AT**C C**

CC**G** GGT AG**C** GCT GAG ACA TAT **C**
C**G** GGT AG**C** GCT GAG ACA TAT **C**
G GGT AG**C** GCT GAG ACA TAT **C**
GGT AG**C** GCT GAG ACA TAT **C**
GT AG**C** GCT GAG ACA TAT **C**
T AG**C** GCT GAG ACA TAT **C**
AG**C** GCT GAG ACA TAT **C**
G**C** GCT GAG ACA TAT **C**
C GCT GAG ACA TAT **C**

GCC **G**GG TAG **C**GC TGA GAC ATA T
CC **G**GG TAG **C**GC TGA GAC ATA T
C **G**GG TAG **C**GC TGA GAC ATA T
GGG TAG **C**GC TGA GAC ATA T
GCC **G**GG TAG **C**GC TGA GAC ATA
GCC **G**GG TAG **C**GC TGA GAC AT

GCC **G**GG TAG **C**GC TGA GAC A
GCC **G**GG TAG **C**GC TGA GAC
GCC **G**GG TAG **C**GC TGA GA
GCC **G**GG TAG **C**GC TGA G
GCC **G**GG TAG **C**GC TGA
GCC **G**GG TAG **C**GC TG

AGC C**G**G GTA G**C**G CTG AGA CAT A
GC C**G**G GTA G**C**G CTG AGA CAT A
C C**G**G GTA G**C**G CTG AGA CAT A
C**G**G GTA G**C**G CTG AGA CAT A
GG GTA G**C**G CTG AGA CAT A
AGC C**G**G GTA G**C**G CTG AGA CAT
AGC C**G**G GTA G**C**G CTG AGA CA
AGC C**G**G GTA G**C**G CTG AGA C
AGC C**G**G GTA G**C**G CTG AGA
AGC C**G**G GTA G**C**G CTG AG
AGC C**G**G GTA G**C**G CTG A
AGC C**G**G GTA G**C**G CTG
AGC C**G**G GTA G**C**G CT

CAG CC**G** GGT AG**C** GCT GAG ACA T
AG CC**G** GGT AG**C** GCT GAG ACA T
G CC**G** GGT AG**C** GCT GAG ACA T
CC**G** GGT AG**C** GCT GAG ACA T
C**G** GGT AG**C** GCT GAG ACA T
G GGT AG**C** GCT GAG ACA T
CAG CC**G** GGT AG**C** GCT GAG ACA
CAG CC**G** GGT AG**C** GCT GAG AC
CAG CC**G** GGT AG**C** GCT GAG A
CAG CC**G** GGT AG**C** GCT GAG
CAG CC**G** GGT AG**C** GCT GA
CAG CC**G** GGT AG**C** GCT G
CAG CC**G** GGT AG**C** GCT
CAG CC**G** GGT AG**C** GC

CAG CC**G** GGT AG**C** GCT GAG ACA T
AG CC**G** GGT AG**C** GCT GAG ACA T

G CCG GGT AGC GCT GAG ACA T
CCG GGT AGC GCT GAG ACA T
CG GGT AGC GCT GAG ACA T
CAG CCG GGT AGC GCT GAG ACA
CAG CCG GGT AGC GCT GAG AC
CAG CCG GGT AGC GCT GAG A
CAG CCG GGT AGC GCT GAG
CAG CCG GGT AGC GCT GA
CAG CCG GGT AGC GCT G
CAG CCG GGT AGC GCT
CAG CCG GGT AGC GC

TCA GCC GGG TAG CGC TGA GAC A
CA GCC GGG TAG CGC TGA GAC A
A GCC GGG TAG CGC TGA GAC A
GCC GGG TAG CGC TGA GAC A
CC GGG TAG CGC TGA GAC A
C GGG TAG CGC TGA GAC A
GGG TAG CGC TGA GAC A
TCA GCC GGG TAG CGC TGA GAC
TCA GCC GGG TAG CGC TGA GA
TCA GCC GGG TAG CGC TGA G
TCA GCC GGG TAG CGC TGA
TCA GCC GGG TAG CGC TG
TCA GCC GGG TAG CGC T
TCA GCC GGG TAG CGC
TCA GCC GGG TAG CG

CTC AGC CGG GTA GCG CTG AGA C
TC AGC CGG GTA GCG CTG AGA C
C AGC CGG GTA GCG CTG AGA C
AGC CGG GTA GCG CTG AGA C
GC CGG GTA GCG CTG AGA C
C CGG GTA GCG CTG AGA C
CGG GTA GCG CTG AGA C
GG GTA GCG CTG AGA C
CTC AGC CGG GTA GCG CTG AGA
CTC AGC CGG GTA GCG CTG AG

CTC AGC CGG GTA GCG CTG A
CTC AGC CGG GTA GCG CTG
CTC AGC CGG GTA GCG CT
CTC AGC CGG GTA GCG C
CTC AGC CGG GTA GCG
CTC AGC CGG GTA GC

TCT CAG CCG GGT AGC GCT GAG A
CT CAG CCG GGT AGC GCT GAG A
T CAG CCG GGT AGC GCT GAG A
CAG CCG GGT AGC GCT GAG A
AG CCG GGT AGC GCT GAG A
G CCG GGT AGC GCT GAG A
CCG GGT AGC GCT GAG A
CG GGT AGC GCT GAG A
G GGT AGC GCT GAG A
TCT CAG CCG GGT AGC GCT GAG
TCT CAG CCG GGT AGC GCT GA
TCT CAG CCG GGT AGC GCT G
TCT CAG CCG GGT AGC GCT
TCT CAG CCG GGT AGC GC
TCT CAG CCG GGT AGC G
TCT CAG CCG GGT AGC

CTC TCA GCC GGG TAG CGC TGA G
TC TCA GCC GGG TAG CGC TGA G
C TCA GCC GGG TAG CGC TGA G
TCA GCC GGG TAG CGC TGA G
CA GCC GGG TAG CGC TGA G
A GCC GGG TAG CGC TGA G
GCC GGG TAG CGC TGA G
CC GGG TAG CGC TGA G
C GGG TAG CGC TGA G
GGG TAG CGC TGA G
CTC TCA GCC GGG TAG CGC TGA
CTC TCA GCC GGG TAG CGC TG
CTC TCA GCC GGG TAG CGC T
CTC TCA GCC GGG TAG CGC

CTC TCA GCC GGG TAG CG
CTC TCA GCC GGG TAG C

CCT CTC AGC CGG GTA GCG CTG A
CT CTC AGC CGG GTA GCG CTG A
T CTC AGC CGG GTA GCG CTG A
CTC AGC CGG GTA GCG CTG A
TC AGC CGG GTA GCG CTG A
C AGC CGG GTA GCG CTG A
AGC CGG GTA GCG CTG A
GC CGG GTA GCG CTG A
C CGG GTA GCG CTG A
CCT CTC AGC CGG GTA GCG CTG
CCT CTC AGC CGG GTA GCG CT
CCT CTC AGC CGG GTA GCG C
CCT CTC AGC CGG GTA GCG
CCT CTC AGC CGG GTA GC

GCC TCT CAG CCG GGT AGC GCT G
CC TCT CAG CCG GGT AGC GCT G
C TCT CAG CCG GGT AGC GCT G
TCT CAG CCG GGT AGC GCT G
CT CAG CCG GGT AGC GCT G
T CAG CCG GGT AGC GCT G
TCT CAG CCG GGT AGC GCT G
CT CAG CCG GGT AGC GCT G
T CAG CCG GGT AGC GCT G
CAG CCG GGT AGC GCT G
AG CCG GGT AGC GCT G
G CCG GGT AGC GCT G
GCC TCT CAG CCG GGT AGC GCT
GCC TCT CAG CCG GGT AGC GC
GCC TCT CAG CCG GGT AGC G
GCC TCT CAG CCG GGT AGC
GCC TCT CAG CCG GGT AG
GCC TCT CAG CCG GGT A
GCC TCT CAG CCG GGT
GCC TCT CAG CCG GG

TGC CTC TCA GCC GGG TAG CGC T
GC CTC TCA GCC GGG TAG CGC T
C CTC TCA GCC GGG TAG CGC T
CTC TCA GCC GGG TAG CGC T
TC TCA GCC GGG TAG CGC T
C TCA GCC GGG TAG CGC T
TCA GCC GGG TAG CGC T
CA GCC GGG TAG CGC T
A GCC GGG TAG CGC T
TGC CTC TCA GCC GGG TAG CGC
TGC CTC TCA GCC GGG TAG CG
TGC CTC TCA GCC GGG TAG C
TGC CTC TCA GCC GGG TAG
TGC CTC TCA GCC GGG TA
TGC CTC TCA GCC GGG T
TGC CTC TCA GCC GGG
TGC CTC TCA GCC GG

CTG CCT CTC AGC CGG GTA GCG C
TG CCT CTC AGC CGG GTA GCG C
G CCT CTC AGC CGG GTA GCG C
CCT CTC AGC CGG GTA GCG C
CT CTC AGC CGG GTA GCG C
T CTC AGC CGG GTA GCG C
CTC AGC CGG GTA GCG C
TC AGC CGG GTA GCG C
C AGC CGG GTA GCG C
CTG CCT CTC AGC CGG GTA GCG
CTG CCT CTC AGC CGG GTA GC
CTG CCT CTC AGC CGG GTA G
CTG CCT CTC AGC CGG GTA
CTG CCT CTC AGC CGG GT
CTG CCT CTC AGC CGG G
CTG CCT CTC AGC CGG
CTG CCT CTC AGC CG

ACT GCC TCT CAG CCG GGT AGC G

CT **G**CC **T**CT CAG CC**G** GGT AG**C** G
T **G**CC **T**CT CAG CC**G** GGT AG**C** G
GCC **T**CT CAG CC**G** GGT AG**C** G
CC **T**CT CAG CC**G** GGT AG**C** G
C **T**CT CAG CC**G** GGT AG**C** G
TCT CAG CC**G** GGT AG**C** G
CT CAG CC**G** GGT AG**C** G
T CAG CC**G** GGT AG**C** G
ACT **G**CC **T**CT CAG CC**G** GGT AG**C**
ACT **G**CC **T**CT CAG CC**G** GGT AG
ACT **G**CC **T**CT CAG CC**G** GGT A
ACT **G**CC **T**CT CAG CC**G** GGT
ACT **G**CC **T**CT CAG CC**G** GG
ACT **G**CC **T**CT CAG CC**G** G
ACT **G**CC **T**CT CAG CC**G**

GAC T**G**C C**T**C TCA GCC **G**GG TAG **C**
AC T**G**C C**T**C TCA GCC **G**GG TAG **C**
C T**G**C C**T**C TCA GCC **G**GG TAG **C**
T**G**C C**T**C TCA GCC **G**GG TAG **C**
GC C**T**C TCA GCC **G**GG TAG **C**
C C**T**C TCA GCC **G**GG TAG **C**
C**T**C TCA GCC **G**GG TAG **C**
TC TCA GCC **G**GG TAG **C**
C TCA GCC **G**GG TAG **C**
GAC T**G**C C**T**C TCA GCC **G**GG TAG
GAC T**G**C C**T**C TCA GCC **G**GG TA
GAC T**G**C C**T**C TCA GCC **G**GG T
GAC T**G**C C**T**C TCA GCC **G**GG
GAC T**G**C C**T**C TCA GCC **G**G
GAC T**G**C C**T**C TCA GCC **G**
GAC T**G**C C**T**C TCA GCC
GAC T**G**C C**T**C TCA GC

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 635 to 658 shown in Figure 1 are the following

GGA GAG GAA AAG GAG GCT CTG A

GGA GAG GAA AAG GAG GCT CTG
GGA GAG GAA AAG GAG GCT CT
GGA GAG GAA AAG GAG GCT C
GGA GAG GAA AAG GAG GCT
GGA GAG GAA AAG GAG GC
GGA GAG GAA AAG GAG G
GGA GAG GAA AAG GAG
GGA GAG GAA AAG GA
GGA GAG GAA AAG G

CGG AGA GGA AAA GGA GGC TCT G
CGG AGA GGA AAA GGA GGC TCT
CGG AGA GGA AAA GGA GGC TC
CGG AGA GGA AAA GGA GGC T
CGG AGA GGA AAA GGA GGC
CGG AGA GGA AAA GGA GG
CGG AGA GGA AAA GGA G
CGG AGA GGA AAA GGA
CGG AGA GGA AAA GG

CCG GAG AGG AAA AGG AGG CTC T
CCG GAG AGG AAA AGG AGG CTC
CCG GAG AGG AAA AGG AGG CT
CCG GAG AGG AAA AGG AGG C
CCG GAG AGG AAA AGG AGG
CCG GAG AGG AAA AGG AG
CCG GAG AGG AAA AGG A
CCG GAG AGG AAA AGG
CCG GAG AGG AAA AG

TCC GGA GAG GAA AAG GAG GCT C
TCC GGA GAG GAA AAG GAG GCT
TCC GGA GAG GAA AAG GAG GC
TCC GGA GAG GAA AAG GAG G
TCC GGA GAG GAA AAG GAG
TCC GGA GAG GAA AAG GA
TCC GGA GAG GAA AAG G
TCC GGA GAG GAA AAG

TCC GGA GAG GAA AA

CTC CGG AGA GGA AAA GGA GGC T
CTC CGG AGA GGA AAA GGA GGC
CTC CGG AGA GGA AAA GGA GG
CTC CGG AGA GGA AAA GGA G
CTC CGG AGA GGA AAA GGA
CTC CGG AGA GGA AAA GG
CTC CGG AGA GGA AAA G
CTC CGG AGA GGA AAA
CTC CGG AGA GGA AA

CCT CCG GAG AGG AAA AGG AGG C
CCT CCG GAG AGG AAA AGG AGG
CCT CCG GAG AGG AAA AGG AG
CCT CCG GAG AGG AAA AGG A
CCT CCG GAG AGG AAA AGG
CCT CCG GAG AGG AAA AG
CCT CCG GAG AGG AAA A
CCT CCG GAG AGG AAA
CCT CCG GAG AGG AA

TCC TCC GGA GAG GAA AAG GAG G
TCC TCC GGA GAG GAA AAG GAG
TCC TCC GGA GAG GAA AAG GA
TCC TCC GGA GAG GAA AAG G
TCC TCC GGA GAG GAA AAG
TCC TCC GGA GAG GAA AA
TCC TCC GGA GAG GAA A
TCC TCC GGA GAG GAA
TCC TCC GGA GAG GA

CTC CTC CGG AGA GGA AAA GGA G
CTC CTC CGG AGA GGA AAA GGA
CTC CTC CGG AGA GGA AAA GG
CTC CTC CGG AGA GGA AAA G
CTC CTC CGG AGA GGA AAA
CTC CTC CGG AGA GGA AA

CTC CTC CGG AGA GGA A
CTC CTC CGG AGA GGA
CTC CTC CGG AGA GG

CCT CCT CCG GAG AGG AAA AGG A
CCT CCT CCG GAG AGG AAA AGG
CCT CCT CCG GAG AGG AAA AG
CCT CCT CCG GAG AGG AAA A
CCT CCT CCG GAG AGG AAA
CCT CCT CCG GAG AGG AA
CCT CCT CCG GAG AGG A
CCT CCT CCG GAG AGG
CCT CCT CCG GAG AG

CCC TCC TCC GGA GAG GAA AAG G
CCC TCC TCC GGA GAG GAA AAG
CCC TCC TCC GGA GAG GAA AA
CCC TCC TCC GGA GAG GAA A
CCC TCC TCC GGA GAG GAA
CCC TCC TCC GGA GAG GA
CCC TCC TCC GGA GAG G
CCC TCC TCC GGA GAG
CCC TCC TCC GGA GA
CC TCC TCC GGA GAG GAA AAG G

ACC CTC CTC CGG AGA GGA AAA G
ACC CTC CTC CGG AGA GGA AAA
ACC CTC CTC CGG AGA GGA AA
ACC CTC CTC CGG AGA GGA A
ACC CTC CTC CGG AGA GGA
ACC CTC CTC CGG AGA GG
ACC CTC CTC CGG AGA G
ACC CTC CTC CGG AGA
ACC CTC CTC CGG AG
CC CTC CTC CGG AGA GGA AAA G
C CTC CTC CGG AGA GGA AAA G
CTC CTC CGG AGA GGA AAA G
TC CTC CGG AGA GGA AAA G

C CTC CGG AGA GGA AAA G
CTC CGG AGA GGA AAA G
TC CTC CGG AGA GGA AAA G
C CTC CGG AGA GGA AAA G

CAC CCT CCT CCG GAG AGG AAA A
CAC CCT CCT CCG GAG AGG AAA
CAC CCT CCT CCG GAG AGG AA
CAC CCT CCT CCG GAG AGG A
CAC CCT CCT CCG GAG AGG
CAC CCT CCT CCG GAG AG
CAC CCT CCT CCG GAG A
CAC CCT CCT CCG GAG
CAC CCT CCT CCG GA
AC CCT CCT CCG GAG AGG AAA A

TCA CCA CCC TCC TCC GG
CA CCA CCC TCC TCC GG
A CCA CCC TCC TCC GG
CCA CCC TCC TCC GG
TCA CCA CCC TCC TCC G
TCA CCA CCC TCC TCC
TCA CCA CCC TCC TC

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 761 to 787 shown in Figure 2 are the following

ATG CGC GTG TGG GTC GC
TG CGC GTG TGG GTC GC
G CGC GTG TGG GTC GC
CGC GTG TGG GTC GC
ATG CGC GTG TGG GTC G
ATG CGC GTG TGG GTC
ATG CGC GTG TGG GT

CGC GTA TGC GCG TGT GGG TCG C
GC GTA TGC GCG TGT GGG TCG C
C GTA TGC GCG TGT GGG TCG C

GTA TGC GCG TGT GGG TCG C
TA TGC GCG TGT GGG TCG C
A TGC GCG TGT GGG TCG C
TGC GCG TGT GGG TCG C
GC GCG TGT GGG TCG C
C GCG TGT GGG TCG C
CGC GTA TGC GCG TGT GGG TCG
CGC GTA TGC GCG TGT GGG TC
CGC GTA TGC GCG TGT GGG T
CGC GTA TGC GCG TGT GGG
CGC GTA TGC GCG TGT GG
CGC GTA TGC GCG TGT G
CGC GTA TGC GCG TGT
CGC GTA TGC GCG TG

GCG CGT ATG CGC GTG TGG GTC G
CG CGT ATG CGC GTG TGG GTC G
G CGT ATG CGC GTG TGG GTC G
CGT ATG CGC GTG TGG GTC G
GT ATG CGC GTG TGG GTC G
T ATG CGC GTG TGG GTC G
ATG CGC GTG TGG GTC G
TG CGC GTG TGG GTC G
G CGC GTG TGG GTC G
GCG CGT ATG CGC GTG TGG GTC
GCG CGT ATG CGC GTG TGG GT
GCG CGT ATG CGC GTG TGG G
GCG CGT ATG CGC GTG TGG
GCG CGT ATG CGC GTG TG
GCG CGT ATG CGC GTG T
GCG CGT ATG CGC GTG
GCG CGT ATG CGC GT

CGC GCG TAT GCG CGT GTG GGT C
GC GCG TAT GCG CGT GTG GGT C
C GCG TAT GCG CGT GTG GGT C
GCG TAT GCG CGT GTG GGT C
CG TAT GCG CGT GTG GGT C

G TAT GCG CGT GTG GGT C
TAT GCG CGT GTG GGT C
AT GCG CGT GTG GGT C
T GCG CGT GTG GGT C
CGC GCG TAT GCG CGT GTG GGT
CGC GCG TAT GCG CGT GTG GG
CGC GCG TAT GCG CGT GTG G
CGC GCG TAT GCG CGT GTG
CGC GCG TAT GCG CGT GT
CGC GCG TAT GCG CGT G
CGC GCG TAT GCG CGT
CGC GCG TAT GCG CG

ACG CGC GTA TGC GCG TGT GGG T
CG CGC GTA TGC GCG TGT GGG T
G CGC GTA TGC GCG TGT GGG T
CGC GTA TGC GCG TGT GGG T
GC GTA TGC GCG TGT GGG T
C GTA TGC GCG TGT GGG T
GTA TGC GCG TGT GGG T
TA TGC GCG TGT GGG T
A TGC GCG TGT GGG T
ACG CGC GTA TGC GCG TGT GGG
ACG CGC GTA TGC GCG TGT GG
ACG CGC GTA TGC GCG TGT G
ACG CGC GTA TGC GCG TGT
ACG CGC GTA TGC GCG TG
ACG CGC GTA TGC GCG T
ACG CGC GTA TGC GCG
ACG CGC GTA TGC GC

CAC GCG CGT ATG CGC GTG TGG G
AC GCG CGT ATG CGC GTG TGG G
C GCG CGT ATG CGC GTG TGG G
GCG CGT ATG CGC GTG TGG G
CG CGT ATG CGC GTG TGG G
G CGT ATG CGC GTG TGG G
CGT ATG CGC GTG TGG G

GT ATG CGC GTG TGG G
T ATG CGC GTG TGG G
CAC GCG CGT ATG CGC GTG TGG
CAC GCG CGT ATG CGC GTG TG
CAC GCG CGT ATG CGC GTG T
CAC GCG CGT ATG CGC GTG
CAC GCG CGT ATG CGC GT
CAC GCG CGT ATG CGC G
CAC GCG CGT ATG CGC
CAC GCG CGT ATG CG

ACA CGC GCG TAT GCG CGT GTG G
CA CGC GCG TAT GCG CGT GTG G
A CGC GCG TAT GCG CGT GTG G
CGC GCG TAT GCG CGT GTG G
GC GCG TAT GCG CGT GTG G
C GCG TAT GCG CGT GTG G
GCG TAT GCG CGT GTG G
CG TAT GCG CGT GTG G
G TAT GCG CGT GTG G
ACA CGC GCG TAT GCG CGT GTG
ACA CGC GCG TAT GCG CGT GT
ACA CGC GCG TAT GCG CGT G
ACA CGC GCG TAT GCG CGT
ACA CGC GCG TAT GCG CG
ACA CGC GCG TAT GCG C
ACA CGC GCG TAT GCG
ACA CGC GCG TAT GC

CAC ACG CGC GTA TGC GCG TGT G
AC ACG CGC GTA TGC GCG TGT G
C ACG CGC GTA TGC GCG TGT G
ACG CGC GTA TGC GCG TGT G
CG CGC GTA TGC GCG TGT G
G CGC GTA TGC GCG TGT G
CGC GTA TGC GCG TGT G
GC GTA TGC GCG TGT G
C GTA TGC GCG TGT G

CAC ACG CGC GTA TGC GCG TGT
CAC ACG CGC GTA TGC GCG TG
CAC ACG CGC GTA TGC GCG T
CAC ACG CGC GTA TGC GCG
CAC ACG CGC GTA TGC GC
CAC ACG CGC GTA TGC G
CAC ACG CGC GTA TGC
CAC ACG CGC GTA TG

TCA CAC GCG CGT ATG CGC GTG T
CA CAC GCG CGT ATG CGC GTG T
A CAC GCG CGT ATG CGC GTG T
CAC GCG CGT ATG CGC GTG T
AC GCG CGT ATG CGC GTG T
C GCG CGT ATG CGC GTG T
GCG CGT ATG CGC GTG T
CG CGT ATG CGC GTG T
G CGT ATG CGC GTG T
TCA CAC GCG CGT ATG CGC GTG
TCA CAC GCG CGT ATG CGC GT
TCA CAC GCG CGT ATG CGC G
TCA CAC GCG CGT ATG CGC
TCA CAC GCG CGT ATG CG
TCA CAC GCG CGT ATG C
TCA CAC GCG CGT ATG
TCA CAC GCG CGT AT

TTC ACA CGC GCG TAT GCG CGT G
TC ACA CGC GCG TAT GCG CGT G
C ACA CGC GCG TAT GCG CGT G
ACA CGC GCG TAT GCG CGT G
CA CGC GCG TAT GCG CGT G
A CGC GCG TAT GCG CGT G
CGC GCG TAT GCG CGT G
GC GCG TAT GCG CGT G
C GCG TAT GCG CGT G
TTC ACA CGC GCG TAT GCG CGT
TTC ACA CGC GCG TAT GCG CG

TTC ACA CGC GCG TAT GCG C
TTC ACA CGC GCG TAT GCG
TTC ACA CGC GCG TAT GC
TTC ACA CGC GCG TAT G
TTC ACA CGC GCG TAT
TTC ACA CGC GCG TA

ATT CAC ACG CGC GTA TGC GCG T
TT CAC ACG CGC GTA TGC GCG T
T CAC ACG CGC GTA TGC GCG T
CAC ACG CGC GTA TGC GCG T
AC ACG CGC GTA TGC GCG T
C ACG CGC GTA TGC GCG T
ACG CGC GTA TGC GCG T
CG CGC GTA TGC GCG T
G CGC GTA TGC GCG T
ATT CAC ACG CGC GTA TGC GCG
ATT CAC ACG CGC GTA TGC GC
ATT CAC ACG CGC GTA TGC G
ATT CAC ACG CGC GTA TGC
ATT CAC ACG CGC GTA TG
ATT CAC ACG CGC GTA T
ATT CAC ACG CGC GTA
ATT CAC ACG CGC GT

TAT TCA CAC GCG CGT ATG CGC G
AT TCA CAC GCG CGT ATG CGC G
T TCA CAC GCG CGT ATG CGC G
TCA CAC GCG CGT ATG CGC G
CA CAC GCG CGT ATG CGC G
A CAC GCG CGT ATG CGC G
CAC GCG CGT ATG CGC G
AC GCG CGT ATG CGC G
C GCG CGT ATG CGC G
TAT TCA CAC GCG CGT ATG CGC
TAT TCA CAC GCG CGT ATG CG
TAT TCA CAC GCG CGT ATG C
TAT TCA CAC GCG CGT ATG

TAT TCA CAC GCG CGT AT
TAT TCA CAC GCG CGT A
TAT TCA CAC GCG CGT
TAT TCA CAC GCG CG

CTA TTC ACA CGC GCG TAT GCG C
TA TTC ACA CGC GCG TAT GCG C
A TTC ACA CGC GCG TAT GCG C
TTC ACA CGC GCG TAT GCG C
TC ACA CGC GCG TAT GCG C
C ACA CGC GCG TAT GCG C
ACA CGC GCG TAT GCG C
CA CGC GCG TAT GCG C
A CGC GCG TAT GCG C
CTA TTC ACA CGC GCG TAT GCG
CTA TTC ACA CGC GCG TAT GC
CTA TTC ACA CGC GCG TAT G
CTA TTC ACA CGC GCG TAT
CTA TTC ACA CGC GCG TA
CTA TTC ACA CGC GCG T
CTA TTC ACA CGC GCG
CTA TTC ACA CGC GC

ACT ATT CAC ACG CGC GTA TGC G
CT ATT CAC ACG CGC GTA TGC G
T ATT CAC ACG CGC GTA TGC G
ATT CAC ACG CGC GTA TGC G
TT CAC ACG CGC GTA TGC G
T CAC ACG CGC GTA TGC G
CAC ACG CGC GTA TGC G
AC ACG CGC GTA TGC G
C ACG CGC GTA TGC G
ACT ATT CAC ACG CGC GTA TGC
ACT ATT CAC ACG CGC GTA TG
ACT ATT CAC ACG CGC GTA T
ACT ATT CAC ACG CGC GTA
ACT ATT CAC ACG CGC GT
ACT ATT CAC ACG CGC G

ACT ATT CAC ACG CGC
ACT ATT CAC ACG CG

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 1306 to 1322 shown in Figure 3 are the following

ACA CCC ACC ACA AGG TGG ATG T
CA CCC ACC ACA AGG TGG ATG T
A CCC ACC ACA AGG TGG ATG T
CCC ACC ACA AGG TGG ATG T
CC ACC ACA AGG TGG ATG T
C ACC ACA AGG TGG ATG T
ACC ACA AGG TGG ATG T
CC ACA AGG TGG ATG T
C ACA AGG TGG ATG T
ACA CCC ACC ACA AGG TGG ATG
ACA CCC ACC ACA AGG TGG AT
ACA CCC ACC ACA AGG TGG A
ACA CCC ACC ACA AGG TGG
ACA CCC ACC ACA AGG TG
ACA CCC ACC ACA AGG T
ACA CCC ACC ACA AGG
ACA CCC ACC ACA AG

CAC ACC CAC CAC AAG GTG GAT G
AC ACC CAC CAC AAG GTG GAT G
C ACC CAC CAC AAG GTG GAT G
ACC CAC CAC AAG GTG GAT G
CC CAC CAC AAG GTG GAT G
C CAC CAC AAG GTG GAT G
CAC CAC AAG GTG GAT G
AC CAC AAG GTG GAT G
C CAC AAG GTG GAT G
CAC ACC CAC CAC AAG GTG GAT
CAC ACC CAC CAC AAG GTG GA
CAC ACC CAC CAC AAG GTG G
CAC ACC CAC CAC AAG GTG
CAC ACC CAC CAC AAG GT

CAC ACC CAC CAC AAG G
CAC ACC CAC CAC AAG
CAC ACC CAC CAC AA

CCA CAC CCA CCA CAA GGT GGA T
CA CAC CCA CCA CAA GGT GGA T
A CAC CCA CCA CAA GGT GGA T
CAC CCA CCA CAA GGT GGA T
AC CCA CCA CAA GGT GGA T
C CCA CCA CAA GGT GGA T
CCA CCA CAA GGT GGA T
CA CCA CAA GGT GGA T
A CCA CAA GGT GGA T
CCA CAC CCA CCA CAA GGT GGA
CCA CAC CCA CCA CAA GGT GG
CCA CAC CCA CCA CAA GGT G
CCA CAC CCA CCA CAA GGT
CCA CAC CCA CCA CAA GG
CCA CAC CCA CCA CAA G
CCA CAC CCA CCA CAA
CCA CAC CCA CCA CA

CCC ACA CCC ACC ACA AGG TGG A
CC ACA CCC ACC ACA AGG TGG A
C ACA CCC ACC ACA AGG TGG A
ACA CCC ACC ACA AGG TGG A
CA CCC ACC ACA AGG TGG A
A CCC ACC ACA AGG TGG A
CCC ACC ACA AGG TGG A
CC ACC ACA AGG TGG A
C ACC ACA AGG TGG A
CCC ACA CCC ACC ACA AGG TGG
CCC ACA CCC ACC ACA AGG TG
CCC ACA CCC ACC ACA AGG T
CCC ACA CCC ACC ACA AGG
CCC ACA CCC ACC ACA AG
CCC ACA CCC ACC ACA A
CCC ACA CCC ACC ACA

CCC ACA **C**CC ACC **AC**

ACC CAC A**C**C CAC C**AC** AAG **G**TG **G**
CC CAC A**C**C CAC C**AC** AAG **G**TG **G**
C CAC A**C**C CAC C**AC** AAG **G**TG **G**
CAC A**C**C CAC C**AC** AAG **G**TG **G**
AC A**C**C CAC C**AC** AAG **G**TG **G**
C A**C**C CAC C**AC** AAG **G**TG **G**
A**C**C CAC C**AC** AAG **G**TG **G**
CC CAC C**AC** AAG **G**TG **G**
C CAC C**AC** AAG **G**TG **G**
ACC CAC A**C**C CAC C**AC** AAG **G**TG
ACC CAC A**C**C CAC C**AC** AAG **G**T
ACC CAC A**C**C CAC C**AC** AAG **G**
ACC CAC A**C**C CAC C**AC** AAG
ACC CAC A**C**C CAC C**AC** AA
ACC CAC A**C**C CAC C**AC** A
ACC CAC A**C**C CAC C**AC**
ACC CAC A**C**C CAC C**A**

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 1621 to 1631 shown in Figure 3 are the following

GCC CC**A** **GAA** CTC CAC **A**CC CCC G
CC CC**A** **GAA** CTC CAC **A**CC CCC G
C CC**A** **GAA** CTC CAC **A**CC CCC G
CC**A** **GAA** CTC CAC **A**CC CCC G
C**A** **GAA** CTC CAC **A**CC CCC G
A **GAA** CTC CAC **A**CC CCC G
GAA CTC CAC **A**CC CCC G
AA CTC CAC **A**CC CCC G
A CTC CAC **A**CC CCC G
GCC CC**A** **GAA** CTC CAC **A**CC CCC
GCC CC**A** **GAA** CTC CAC **A**CC CC
GCC CC**A** **GAA** CTC CAC **A**CC C
GCC CC**A** **GAA** CTC CAC **A**CC
GCC CC**A** **GAA** CTC CAC **AC**
GCC CC**A** **GAA** CTC CAC **A**

GCC CCA GAA CTC CAC
GCC CCA GAA CTC CA

CC CCA GAA CTC CAC ACC CCC
C CCA GAA CTC CAC ACC CC
CCA GAA CTC CAC ACC C
CCA GAA CTC CAC ACC
CA GAA CTC CAC ACC

ACT CCA CAC CCC CGA

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 2401 to 2418 shown in Figure 4 are the following

CCG CCC CAA CTG GCG TCG AGG T
CG CCC CAA CTG GCG TCG AGG T
G CCC CAA CTG GCG TCG AGG T
CCC CAA CTG GCG TCG AGG T
CC CAA CTG GCG TCG AGG T
C CAA CTG GCG TCG AGG T
CAA CTG GCG TCG AGG T
AA CTG GCG TCG AGG T
A CTG GCG TCG AGG T
CCG CCC CAA CTG GCG TCG AGG
CCG CCC CAA CTG GCG TCG AG
CCG CCC CAA CTG GCG TCG A
CCG CCC CAA CTG GCG TCG
CCG CCC CAA CTG GCG TC
CCG CCC CAA CTG GCG T
CCG CCC CAA CTG GCG
CCG CCC CAA CTG GC

TCC GCC CCA ACT GGC GTC GAG G
CC GCC CCA ACT GGC GTC GAG G
C GCC CCA ACT GGC GTC GAG G
GCC CCA ACT GGC GTC GAG G
CC CCA ACT GGC GTC GAG G
C CCA ACT GGC GTC GAG G

CCA ACT GGC GTC **GAG** G
CA ACT GGC GTC **GAG** G
A ACT GGC GTC **GAG** G
TCC **GCC** CCA ACT GGC GTC **GAG**
TCC **GCC** CCA ACT GGC GTC **GA**
TCC **GCC** CCA ACT GGC GTC **G**
TCC **GCC** CCA ACT GGC GTC
TCC **GCC** CCA ACT GGC GT
TCC **GCC** CCA ACT GGC G
TCC **GCC** CCA ACT GGC
TCC **GCC** CCA ACT GG

CTC CGC **CCC** AAC TGG CGT C**GA G**
TC CGC **CCC** AAC TGG CGT C**GA G**
C CGC **CCC** AAC TGG CGT C**GA G**
CGC **CCC** AAC TGG CGT C**GA G**
GC **CCC** AAC TGG CGT C**GA G**
C **CCC** AAC TGG CGT C**GA G**
CCC AAC TGG CGT C**GA G**
CC AAC TGG CGT C**GA G**
C AAC TGG CGT C**GA G**
CTC CGC **CCC** AAC TGG CGT C**GA**
CTC CGC **CCC** AAC TGG CGT C**G**
CTC CGC **CCC** AAC TGG CGT C
CTC CGC **CCC** AAC TGG CGT
CTC CGC **CCC** AAC TGG CG
CTC CGC **CCC** AAC TGG C
CTC CGC **CCC** AAC TGG
CTC CGC **CCC** AAC TG

GCC CCA ACT GGC GTC

ACT CC**G** CCC **C**AA CTG GCG TC**G A**
CT CC**G** CCC **C**AA CTG GCG TC**G A**
T CC**G** CCC **C**AA CTG GCG TC**G A**
CC**G** CCC **C**AA CTG GCG TC**G A**
C**G** CCC **C**AA CTG GCG TC**G A**
G CCC **C**AA CTG GCG TC**G A**

CCC **C**AA CTG GCG TC**G A**
CC **C**AA CTG GCG TC**G A**
C **C**AA CTG GCG TC**G A**
ACT CC**G** C**C**C **C**AA CTG GCG TC**G**
ACT CC**G** C**C**C **C**AA CTG GCG TC
ACT CC**G** C**C**C **C**AA CTG GCG T
ACT CC**G** C**C**C **C**AA CTG GCG
ACT CC**G** C**C**C **C**AA CTG GC
ACT CC**G** C**C**C **C**AA CTG G
ACT CC**G** C**C**C **C**AA CTG
ACT CC**G** C**C**C **C**AA CT

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 2455 to 2486 shown in Figure 4 are the following

CTA AAC CCG ATT **C**AG GGT TCG A
CTA AAC CCG ATT **C**AG GGT TCG
CTA AAC CCG ATT **C**AG GGT TC
CTA AAC CCG ATT **C**AG GGT T
CTA AAC CCG ATT **C**AG GGT
CTA AAC CCG ATT **C**AG GG
CTA AAC CCG ATT **C**AG G
CTA AAC CCG ATT **C**AG
CTA AAC CCG ATT **C**A

CAG GGT TCG A**G**G TTA G**A**T G**C**C C
AG GGT TCG A**G**G TTA G**A**T G**C**C C
G GGT TCG A**G**G TTA G**A**T G**C**C C
GGT TCG A**G**G TTA G**A**T G**C**C C
GT TCG A**G**G TTA G**A**T G**C**C C
T TCG A**G**G TTA G**A**T G**C**C C
TCG A**G**G TTA G**A**T G**C**C C
CG A**G**G TTA G**A**T G**C**C C
G A**G**G TTA G**A**T G**C**C C
CAG GGT TCG A**G**G TTA G**A**T G**C**C
CAG GGT TCG A**G**G TTA G**A**T G**C**
CAG GGT TCG A**G**G TTA G**A**T G
CAG GGT TCG A**G**G TTA G**A**T

CAG GGT TCG AGG TTA GA

TCA GGG TTC GAG GTT AGA TGC C
CA GGG TTC GAG GTT AGA TGC C
A GGG TTC GAG GTT AGA TGC C
GGG TTC GAG GTT AGA TGC C
GG TTC GAG GTT AGA TGC C
G TTC GAG GTT AGA TGC C
TTC GAG GTT AGA TGC C
TC GAG GTT AGA TGC C
C GAG GTT AGA TGC C
TTC GAG GTT AGA TGC
TCA GGG TTC GAG GTT AGA TGC
TCA GGG TTC GAG GTT AGA TG
TCA GGG TTC GAG GTT AGA T
TCA GGG TTC GAG GTT AGA

TTC AGG GTT CGA GGT TAG ATG C
TC AGG GTT CGA GGT TAG ATG C
C AGG GTT CGA GGT TAG ATG C
AGG GTT CGA GGT TAG ATG C
GG GTT CGA GGT TAG ATG C
G GTT CGA GGT TAG ATG C
GTT CGA GGT TAG ATG C
TT CGA GGT TAG ATG C
T CGA GGT TAG ATG C
TTC AGG GTT CGA GGT TAG ATG
TTC AGG GTT CGA GGT TAG AT
TTC AGG GTT CGA GGT TAG A

CTG TCC CTA AAC CCG
GTC CCT AAA CCC GAT

Examples of nucleobase sequences of peptide nucleic acid probes wherein the Qs of adjacent moieties are selected so as to be complementary to nucleobases within Position 3094 to 3103 shown in Figure 4 are the following

CAG GTC TGA CCT ATT GAA CCC G
AG GTC TGA CCT ATT GAA CCC G

G GTC TGA CCT ATT GAA CCC G
GTC TGA CCT ATT GAA CCC G
TC TGA CCT ATT GAA CCC G
C TGA CCT ATT GAA CCC G
TGA CCT ATT GAA CCC G
GA CCT ATT GAA CCC G
A CCT ATT GAA CCC G
CAG GTC TGA CCT ATT GAA CCC
CAG GTC TGA CCT ATT GAA CC
CAG GTC TGA CCT ATT GAA C
CAG GTC TGA CCT ATT GAA
CAG GTC TGA CCT ATT GA
CAG GTC TGA CCT ATT G
CAG GTC TGA CCT ATT
CAG GTC TGA CCT AT

AG GTC TGA CCT ATT GAA CCC
G GTC TGA CCT ATT GAA CC
GTC TGA CCT ATT GAA C
TC TGA CCT ATT GAA

EXAMPLES

EXAMPLE 1

In situ hybridisation to fixed bacterial cells

To test the ability of the peptide nucleic acid probes to detect MTC and not MAC or Neisseria gonorrhoeae, fluorescence *in situ* hybridisation (FISH) was performed on fixed bacterial cells using fluorescein labelled probes as shown below. It was shown that these probes did not hybridise to M. avium, M. intracellulare, or N. gonorrhoeae.

Preparation of bacterial slides

M. bovis BCG (Statens Seruminstitut, Denmark, Catalogue number 2645), M. avium (Statens Seruminstitut, Denmark, Laboratory number 3716 (E37978)), and M. intracellulare (Statens Seruminstitut, Laboratory number 3717 (E39562)) were grown in Dubos medium (Statens Seruminstitut, Denmark) or on Löwenstein-Jensen medium (Statens Seruminstitut, Denmark) at 37 °C. N. gonorrhoeae was grown on chocolate agar at 37 °C with additional 5% CO_2.

Bacterial smears were prepared on test slides according to standard procedures. The smears were air-dried followed by flame fixation.

FISH on bacterial slides

The following procedure was performed.

1. The slide is immersed in 80% ethanol for 15 minutes, subsequently rinsed with water and air-dried.
2. The bacterial slide is covered with a hybridisation solution containing the probe in question at a concentration of 250 nM.
3. The slide is incubated in a humid incubation chamber at 45 °C for 90 minutes.
4. The slide is washed 25 minutes in TBS-buffer, pH 10 at 45 °C, followed by 30 seconds in water.
5. The slide is dried and mounted (DAKO Fluorescence Mounting Medium or equivalent).

The following hybridisation solutions was used:

Hybridisation solution	10 mM NaCl 10% Dextran sulphate 30% formamide 0.1% Triton X-100® 50 mM Tris-HCl, pH 7.6 50 mM EDTA 0.1% sodium pyrophosphate 0.2% polyvinylpyrrolidone 0.2% Ficol
TBS buffer	10 mM sodium phosphate, pH 10 145 mM NaCl

All solutions are made RNase free following standard procedures.

The following peptide nucleic acid probe was used

Lys(Flu)-Lys(Flu)-TTC GAG GTT AGA TGC-NH_2	OK 306
Lys(Flu)-Lys(Flu)-ACT CCA CAC CCC CGA-NH_2	OK 309

wherein Flu denotes a fluorescein isothiocyanate label or a 5-(and 6)-carboxyfluorescein label, and Lys(Flu)-Lys(Flu) denotes a peptide label ("zipper") with two Flu labels attached. The

results are shown in Table 1.

TABLE 1

Probe OK 306	FISH
M. bovis BCG	positive
M. avium	negative
M. intracellulare	negative
N. gonorrhoeae	negative

Probe OK 309	FISH
M. bovis BCG	positive
M. avium	negative
M. intracellulare	negative
N. gonorrhoeae	negative

EXAMPLE 2

Test in dot blots

To further test the ability of the peptide nucleic acid probes to detect MTC and not MAC or E. coli, dot blot tests were carried out.

M. bovis BCG (Statens Seruminstitut Catalogue number 2645) and M. intracellulare (Statens Seruminstitut, Denmark Laboratory number 3713 (E39562)) were grown in Dubos medium (Statens Seruminstitut, Denmark) or on Löwenstein-Jensen medium (Statens Seruminstitut, Denmark) at 37 °C.

RNA was isolated from the bacterial cells by use of TRI-reagent (Sigma) following manufacture's directions. E. coli rRNA was purchased from Boehringer Mannheim, Germany.

The following nucleic acid probes were used.

Lys(Flu)-Lys(Flu)-CTG TCC CTA AAC CCG-NH_2 OK 305

Lys(Flu)-Lys(Flu)-GTC CCT AAA CCC GAT-NH_2 OK 307

Lys(Flu)-Lys(Flu)-ACT CCA CAC CCC CGA-NH_2 OK 309

wherein Flu denotes a fluorescein isothiocyanate label or a 5-(and 6)-carboxyfluorescein label, and Lys(Flu)-Lys(Flu) denotes a peptide label ("zipper") consisting of 2 amino acids, respectively, with two Flu labels attached.

Preparation of dot blots

The following buffers were used:

20 × SSPE buffer	3 M NaCl
	0.2 M PO_4^{3-}
	0.02 M EDTA
	pH 7.4
TST buffer	0.05 M Tris/HCl
	0.5 M NaCl
	0.5% Tween 20®
	pH 9.0

200 ng M. bovis RNA, M. intracellulare RNA and E. coli rRNA were dotted onto membranes (Schleich & Schuel, NY 13 N), and the membranes were dried and fixed under UV light for 2 minutes. Each of the probes (70 nM probe in hybridisation solution (hybridisation solution without Triton X-100® and with the exception that formamide was substituted with 50% glycerol)) were added to the membrane. Hybridisation was continued for 1.5 hours at 55 °C. The membranes were rinsed 2 times for 15 minutes in 2 × SSPE buffer containing 0.1% SDS at ambient temperature, and subsequently 2 times for 15 minutes in 0.1 × SSPE buffer containing 0.1% SDS at 55 °C or at 65 °C (see Table 2). The membrane was blocked with 0.5% casein dissolved in 0.05M Tris/HCl and 0.5 M NaCl with pH 9.0. Thereafter, the membranes were incubated for 1 hour with rabbit-anti FITC antibody labelled with AP (DAKO K0046 vial A) diluted 1:2000 in 0.5% casein dissolved in 0.05M Tris/HCl and 0.5 M NaCl with pH 9.0. After incubation, the membranes were washed 3 times 5 minutes with TST at ambient temperature. Bound probes were visualised following standard procedures using BCIP/NBT, and the visualisation was stopped by incubation for 10 minutes with 10 mM EDTA. The blot was dried at 50 °C.

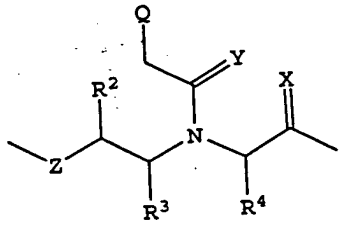
The results are given in Table 2 below.

TABLE 2

	E. coli rRNA		M. bovis BCG RNA		M. intracellulare RNA	
Probe	55 °C	65 °C	55 °C	65 °C	55 °C	65 °C
OK 305	negative	negative	positive	positive	negative	weak
OK 307	negative	negative	positive	positive	negative	weak
OK 309	negative	negative	positive	positive	negative	weak

CLAIMS

1. A peptide nucleic acid probe for detecting mycobacteria of the Mycobacterium tuberculosis Complex (MTC) in a sample, which probe comprises from 10 to 30 polymerised moieties of formula (I)



(I)

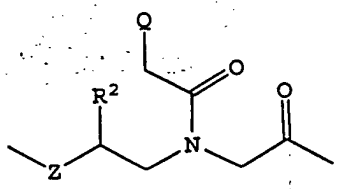
wherein each X and Y independently designate O or S,
each Z independently designates O, S, NR^1, or $C(R^1)_2$, wherein each R^1 independently designate H, C_{1-6} alkyl, C_{1-6} alkenyl, C_{1-6} alkynyl,
each R^2, R^3 and R^4 designate independently H, the side chain of a naturally occurring amino acid, the side chain of a non-naturally occurring nucleobase, C_{1-4} alkyl, C_{1-4} alkenyl or C_{1-4} alkynyl, or a functional group, each Q independently designates a naturally occurring nucleobase, a non-naturally occurring nucleobase, an intercalator, a nucleobase-binding group, a label or H,

with the proviso that the Qs of adjacent moieties are selected so as to form a sequence of which a subsequence includes at least one nucleobase complementary to a nucleobase of M. tuberculosis 23S rRNA that differs from the corresponding nucleobase of M. avium located within the following domains

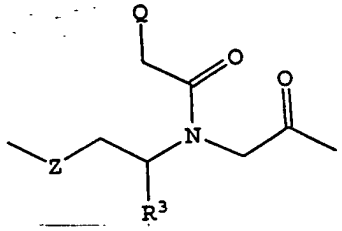
Position 326 to Position 359 in Figure 1, or
Position 635 to Position 658 in Figure 1, or
Position 761 to Position 787 in Figure 2, or
Position 1306 to Position 1322 in Figure 3, or
Position 1621 to Position 1631 in Figure 3, or
Position 2401 to Position 2418 in Figure 4, or
Position 2455 to Position 2486 in Figure 4, or
Position 3094 to Position 3103 in Figure 4,

and further with the proviso that the probe comprising such subsequence is able to form hybrids with target sequences in 23S rRNA of said mycobacteria.

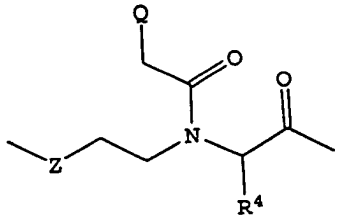
2. A peptide nucleic acid probe according to claim 1 of formula (II), (III), or (IV)



(II)



(III)



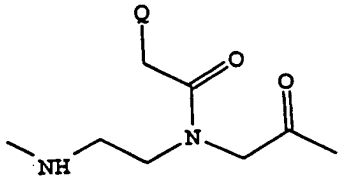
(IV)

wherein Z, R^2, R^3, and R^4, and Q is as defined in claim 1.

3. A peptide nucleic acid probe according to claim 1 or 2, wherein Z is NH, NCH_3 or O, each R^2, R^3 and R^4 independently designate the side chain of a naturally occurring nucleobase, the side chain of a non-naturally occurring nucleobase, or C_{1-4} alkyl, and each Q is a naturally occurring nucleobase or a non-naturally occurring nucleobase with the provisos defined in claim 1.

4. A peptide nucleic acid probe according to anyone of claims 1 to 3, wherein Z is NH or O, and R^2 is H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, and Q is a nucleobase selected from thymine, adenine, cytosine, guanine, uracil, iso-C, iso-G and 2,6-diaminopurine with the provisos defined in claim 1.

5. A peptide nucleic acid probe according to anyone of claims 1 to 4 of formula (V)



(V)

wherein R^4 is H or the side chain of Ala, Asp, Cys, Glu, His, HomoCys, Lys, Orn, Ser or Thr, and Q is as defined in claim 4 with the provisos defined in claim 1.

6. A peptide nucleic acid probe according to anyone of claims 1 to 5 further comprising one or more labels which may be mutually identical or different, and/or one or more linkers which may be mutually identical or different with the provisos defined in claim 1.

7. Method for detecting mycobacteria of the Mycobacterium tuberculosis Complex in a sample comprising

(1) contacting any rRNA optionally present in said sample with one or more peptide nucleic acid probes according to anyone of claims 1 to 6 under conditions, whereby hybrids between said probe(s) and said rRNA are formed, and

(2) observing or measuring said hybridisation, and relating said observation or measurement to the presence of mycobacteria of the Mycobacterium tuberculosis Complex in said sample.

8. Method according to claim 7,
c h a r a c t e r i s e d in that the hybrids are captured on a solid phase before measuring the extent of hybridisation.

9. Method according to claim 7,
c h a r a c t e r i s e d in that a peptide nucleic acid probe according to anyone of claims 1 to 6 are used for capturing the hybrids.

10. A method according to anyone of claims 7 to 9,
c h a r a c t e r i s e d in that a signal amplifying system is used for measuring the resulting hybridisation.

11. Kit for detecting mycobacteria of the Mycobacterium tuberculosis Complex,
c h a r a c t e r i s e d in that said kit comprises at least one peptide nucleic acid probe according to anyone of claims 1 to 6, and a detection system with at least one detecting reagent.

12. Kit according to claim 11,
c h a r a c t e r i s e d in that it further comprises a solid phase capture system.

ABSTRACT

NOVEL PROBES FOR THE DETECTION OF MYCOBACTERIA OF THE MYCOBACTERIUM TUBERCULOSIS COMPLEX

Novel hybridisation assay probes for detecting mycobacteria of the Mycobacterium tuberculosis Complex (MTC) are provided. The probes detect 23S rRNA of MTC. Such probes are capable of detecting the organisms in test samples, e.g. expectorates, sputum, aspirates, urine, blood and tissue sections, food, soil and water.

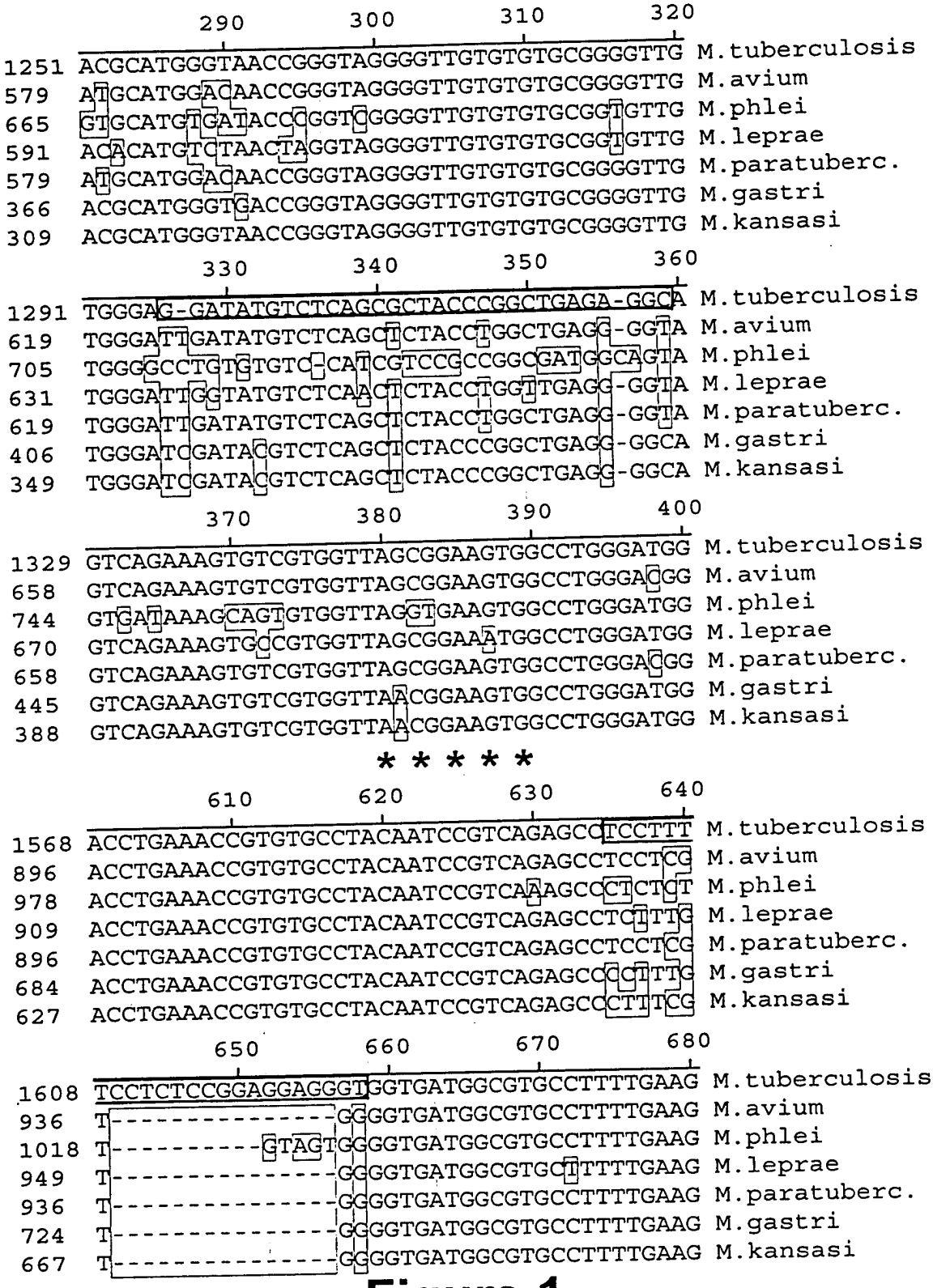

290 300 310 320
1251 ACGCATGGGTAACCGGGTAGGGGTTGTGTGTGCGGGGTTG M.tuberculosis
579 ATGCATGGACAACCGGGTAGGGGTTGTGTGTGCGGGGTTG M.avium
665 GTGCATGTGATACCGGTCGGGGTTGTGTGTGCGGTGTTG M.phlei
591 ACACATGTCTAACTAGGTAGGGGTTGTGTGTGCGGTGTTG M.leprae
579 ATGCATGGACAACCGGGTAGGGGTTGTGTGTGCGGGGTTG M.paratuberc.
366 ACGCATGGGTGACCGGGTAGGGGTTGTGTGTGCGGGGTTG M.gastri
309 ACGCATGGGTAACCGGGTAGGGGTTGTGTGTGCGGGGTTG M.kansasi
330 340 350 360
1291 TGGGAG-GATATGTCTCAGCGCTACCCGGCTGAGA-GGCA M.tuberculosis
619 TGGGATTGATATGTCTCAGCTCTACCTGGCTGAGG-GGTA M.avium
705 TGGGGCCTGTGTGTC-CATCGTCCGCCGGCGATGGCAGTA M.phlei
631 TGGGATTGGTATGTCTCAACTCTACCTGGTTGAGG-GGTA M.leprae
619 TGGGATTGATATGTCTCAGCTCTACCTGGCTGAGG-GGTA M.paratuberc.
406 TGGGATCGATACGTCTCAGCTCTACCCGGCTGAGG-GGCA M.gastri
349 TGGGATCGATACGTCTCAGCTCTACCCGGCTGAGG-GGCA M.kansasi
370 380 390 400
1329 GTCAGAAAGTGTCGTGGTTAGCGGAAGTGGCCTGGGATGG M.tuberculosis
658 GTCAGAAAGTGTCGTGGTTAGCGGAAGTGGCCTGGGACGG M.avium
744 GTGATAAAGCAGTGTGGTTAGGTGAAGTGGCCTGGGATGG M.phlei
670 GTCAGAAAGTGCCGTGGTTAGCGGAAATGGCCTGGGATGG M.leprae
658 GTCAGAAAGTGTCGTGGTTAGCGGAAGTGGCCTGGGACGG M.paratuberc.
445 GTCAGAAAGTGTCGTGGTTAACGGAAGTGGCCTGGGATGG M.gastri
388 GTCAGAAAGTGTCGTGGTTAACGGAAGTGGCCTGGGATGG M.kansasi
* * * * *
610 620 630 640
1568 ACCTGAAACCGTGTGCCTACAATCCGTCAGAGCCTCCTTT M.tuberculosis
896 ACCTGAAACCGTGTGCCTACAATCCGTCAGAGCCTCCTCG M.avium
978 ACCTGAAACCGTGTGCCTACAATCCGTCAAAGCCCTCTCT M.phlei
909 ACCTGAAACCGTGTGCCTACAATCCGTCAGAGCCTCTTTG M.leprae
896 ACCTGAAACCGTGTGCCTACAATCCGTCAGAGCCTCCTCG M.paratuberc.
684 ACCTGAAACCGTGTGCCTACAATCCGTCAGAGCCCCTTTG M.gastri
627 ACCTGAAACCGTGTGCCTACAATCCGTCAGAGCCCTTTCG M.kansasi
650 660 670 680
1608 TCCTCTCCGGAGGAGGGTGGTGATGGCGTGCCTTTTGAAG M.tuberculosis
936 T----------GGGGTGATGGCGTGCCTTTTGAAG M.avium
1018 T----------GTAGTGGGGTGATGGCGTGCCTTTTGAAG M.phlei
949 T----------GGGGTGATGGCGTGCTTTTTGAAG M.leprae
936 T----------GGGGTGATGGCGTGCCTTTTGAAG M.paratuberc.
724 T----------GGGGTGATGGCGTGCCTTTTGAAG M.gastri
667 T----------GGGGTGATGGCGTGCCTTTTGAAG M.kansasi

Figure 1

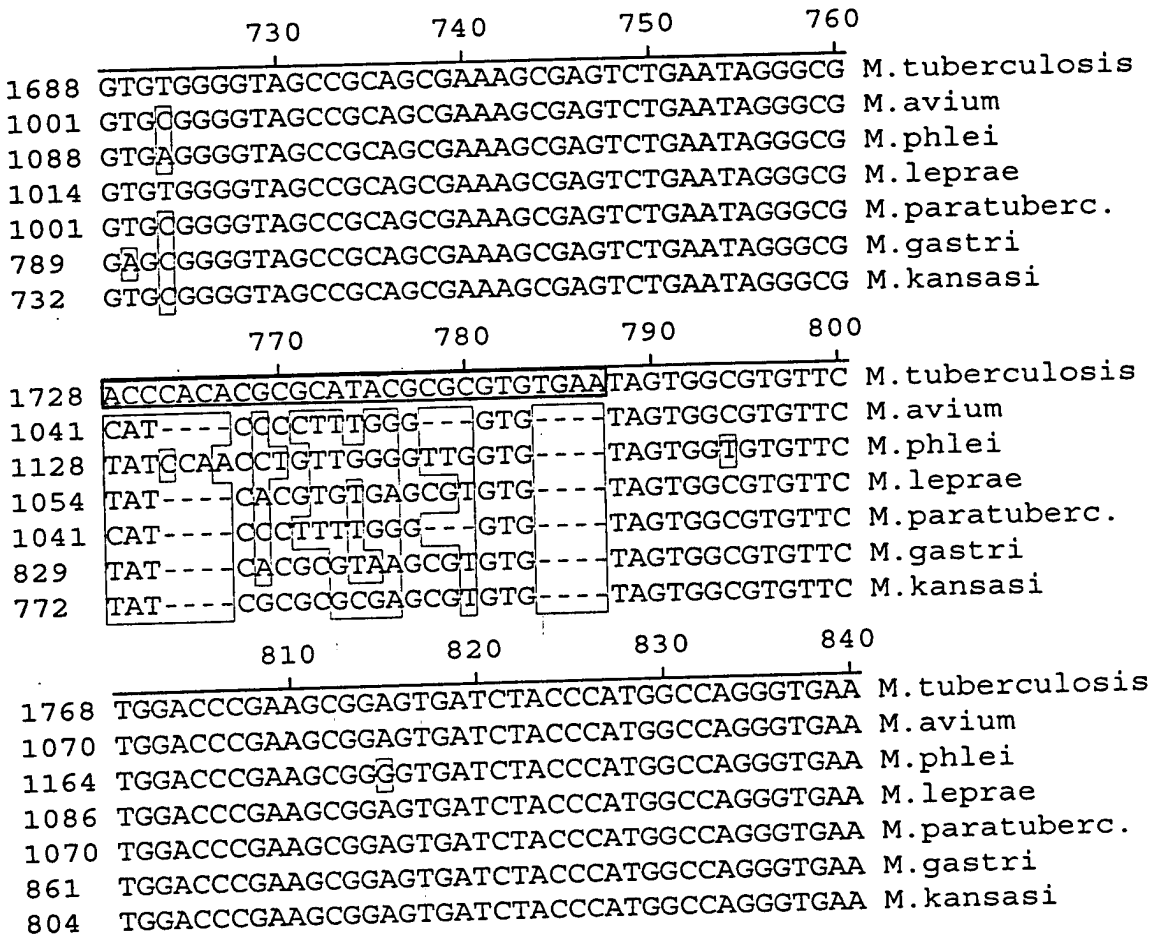

730 740 750 760
1688 GTGTGGGGTAGCCGCAGCGAAAGCGAGTCTGAATAGGGCG M.tuberculosis
1001 GTGCGGGGTAGCCGCAGCGAAAGCGAGTCTGAATAGGGCG M.avium
1088 GTGAGGGGTAGCCGCAGCGAAAGCGAGTCTGAATAGGGCG M.phlei
1014 GTGTGGGGTAGCCGCAGCGAAAGCGAGTCTGAATAGGGCG M.leprae
1001 GTGCGGGGTAGCCGCAGCGAAAGCGAGTCTGAATAGGGCG M.paratuberc.
789 GAGCGGGGTAGCCGCAGCGAAAGCGAGTCTGAATAGGGCG M.gastri
732 GTGCGGGGTAGCCGCAGCGAAAGCGAGTCTGAATAGGGCG M.kansasi
770 780 790 800
1728 ACCCACACGCGCATACGCGCGTGTGAATAGTGGCGTGTTC M.tuberculosis
1041 CAT----CCCCTTTGGG---GTG----TAGTGGCGTGTTC M.avium
1128 TATCCAACCTGTTGGGGTTGGTG----TAGTGGTGTGTTC M.phlei
1054 TAT----CACGTGTGAGCGTGTG----TAGTGGCGTGTTC M.leprae
1041 CAT----CCCTTTTGGG---GTG----TAGTGGCGTGTTC M.paratuberc.
829 TAT----CACGCGTAAGCGTGTG----TAGTGGCGTGTTC M.gastri
772 TAT----CGCGCGCGAGCGTGTG----TAGTGGCGTGTTC M.kansasi
810 820 830 840
1768 TGGACCCGAAGCGGAGTGATCTACCCATGGCCAGGGTGAA M.tuberculosis
1070 TGGACCCGAAGCGGAGTGATCTACCCATGGCCAGGGTGAA M.avium
1164 TGGACCCGAAGCGGGGTGATCTACCCATGGCCAGGGTGAA M.phlei
1086 TGGACCCGAAGCGGAGTGATCTACCCATGGCCAGGGTGAA M.leprae
1070 TGGACCCGAAGCGGAGTGATCTACCCATGGCCAGGGTGAA M.paratuberc.
861 TGGACCCGAAGCGGAGTGATCTACCCATGGCCAGGGTGAA M.gastri
804 TGGACCCGAAGCGGAGTGATCTACCCATGGCCAGGGTGAA M.kansasi

Figure 2

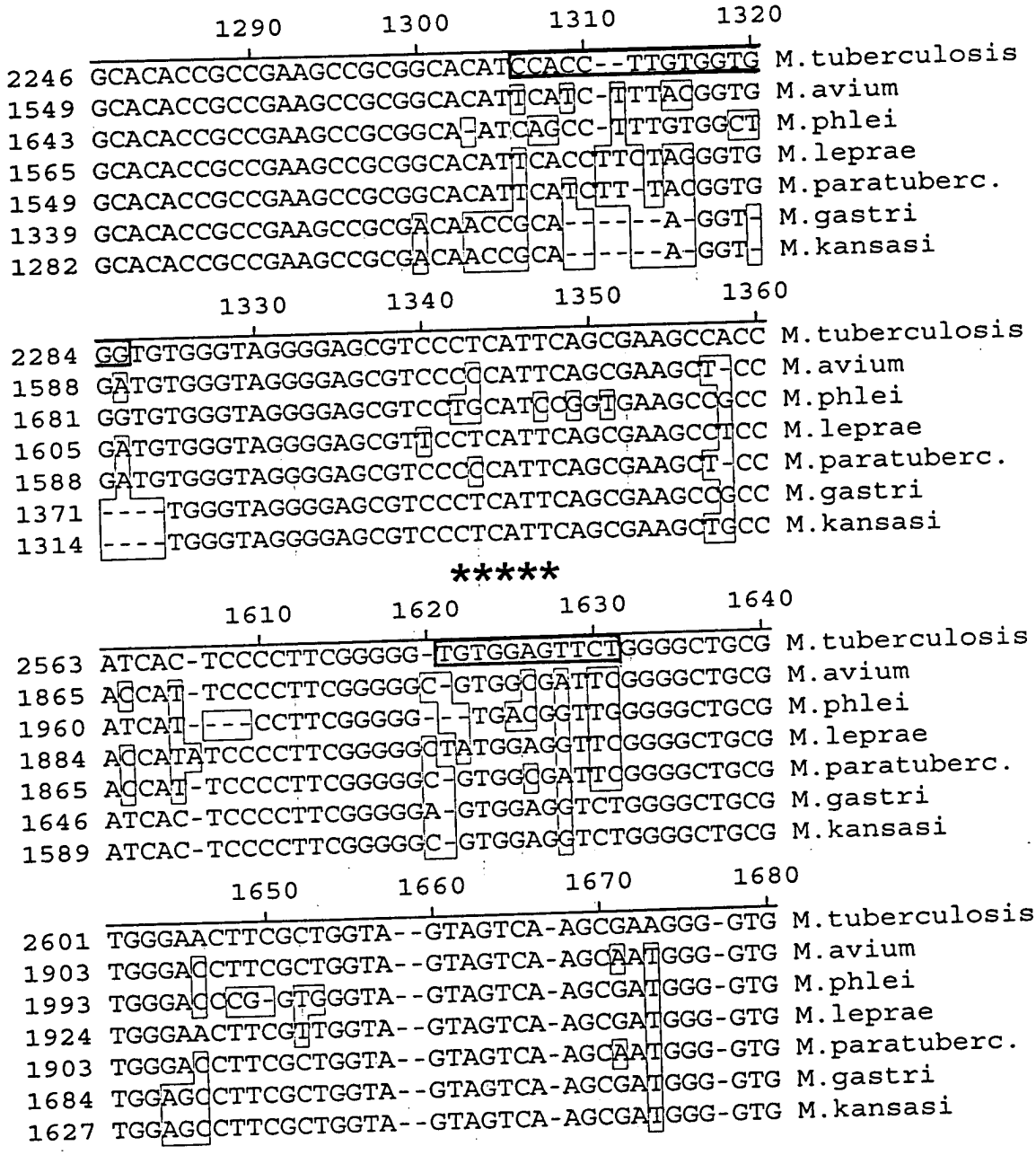

1290 1300 1310 1320
2246 GCACACCGCCGAAGCCGCGGCACATCCACC--TTGTGGTG M.tuberculosis
1549 GCACACCGCCGAAGCCGCGGCACATTCATC-TTTACGGTG M.avium
1643 GCACACCGCCGAAGCCGCGGCA-ATCAGCC-TTTGTGGCT M.phlei
1565 GCACACCGCCGAAGCCGCGGCACATTCACCTTCTAGGGTG M.leprae
1549 GCACACCGCCGAAGCCGCGGCACATTCATCTT-TACGGTG M.paratuberc.
1339 GCACACCGCCGAAGCCGCGACAACCGCA-----A-GGT- M.gastri
1282 GCACACCGCCGAAGCCGCGACAACCGCA-----A-GGT- M.kansasi
1330 1340 1350 1360
2284 GGTGTGGGTAGGGGAGCGTCCCTCATTCAGCGAAGCCACC M.tuberculosis
1588 GATGTGGGTAGGGGAGCGTCCCCCATTCAGCGAAGCT-CC M.avium
1681 GGTGTGGGTAGGGGAGCGTCCTGCATCCGGTGAAGCCGCC M.phlei
1605 GATGTGGGTAGGGGAGCGTTCCTCATTCAGCGAAGCCTCC M.leprae
1588 GATGTGGGTAGGGGAGCGTCCCCCATTCAGCGAAGCT-CC M.paratuberc.
1371 ----TGGGTAGGGGAGCGTCCCTCATTCAGCGAAGCCGCC M.gastri
1314 ----TGGGTAGGGGAGCGTCCCTCATTCAGCGAAGCTGCC M.kansasi
*****
1610 1620 1630 1640
2563 ATCAC-TCCCCTTCGGGGG-TGTGGAGTTCTGGGGCTGCG M.tuberculosis
1865 ACCAT-TCCCCTTCGGGGGC-GTGGCGATTCGGGGCTGCG M.avium
1960 ATCAT-----CCTTCGGGGG---TGACGGTTGGGGGCTGCG M.phlei
1884 ACCATATCCCCTTCGGGGGCTATGGAGGTTCGGGGCTGCG M.leprae
1865 ACCAT-TCCCCTTCGGGGGC-GTGGCGATTCGGGGCTGCG M.paratuberc.
1646 ATCAC-TCCCCTTCGGGGGA-GTGGAGGTCTGGGGCTGCG M.gastri
1589 ATCAC-TCCCCTTCGGGGGC-GTGGAGGTCTGGGGCTGCG M.kansasi
1650 1660 1670 1680
2601 TGGGAACTTCGCTGGTA--GTAGTCA-AGCGAAGGG-GTG M.tuberculosis
1903 TGGGACCTTCGCTGGTA--GTAGTCA-AGCAATGGG-GTG M.avium
1993 TGGGACCCG-GTGGGTA--GTAGTCA-AGCGATGGG-GTG M.phlei
1924 TGGGAACTTCGTTGGTA--GTAGTCA-AGCGATGGG-GTG M.leprae
1903 TGGGACCTTCGCTGGTA--GTAGTCA-AGCAATGGG-GTG M.paratuberc.
1684 TGGAGCCTTCGCTGGTA--GTAGTCA-AGCGATGGG-GTG M.gastri
1627 TGGAGCCTTCGCTGGTA--GTAGTCA-AGCGATGGG-GTG M.kansasi

Figure 3

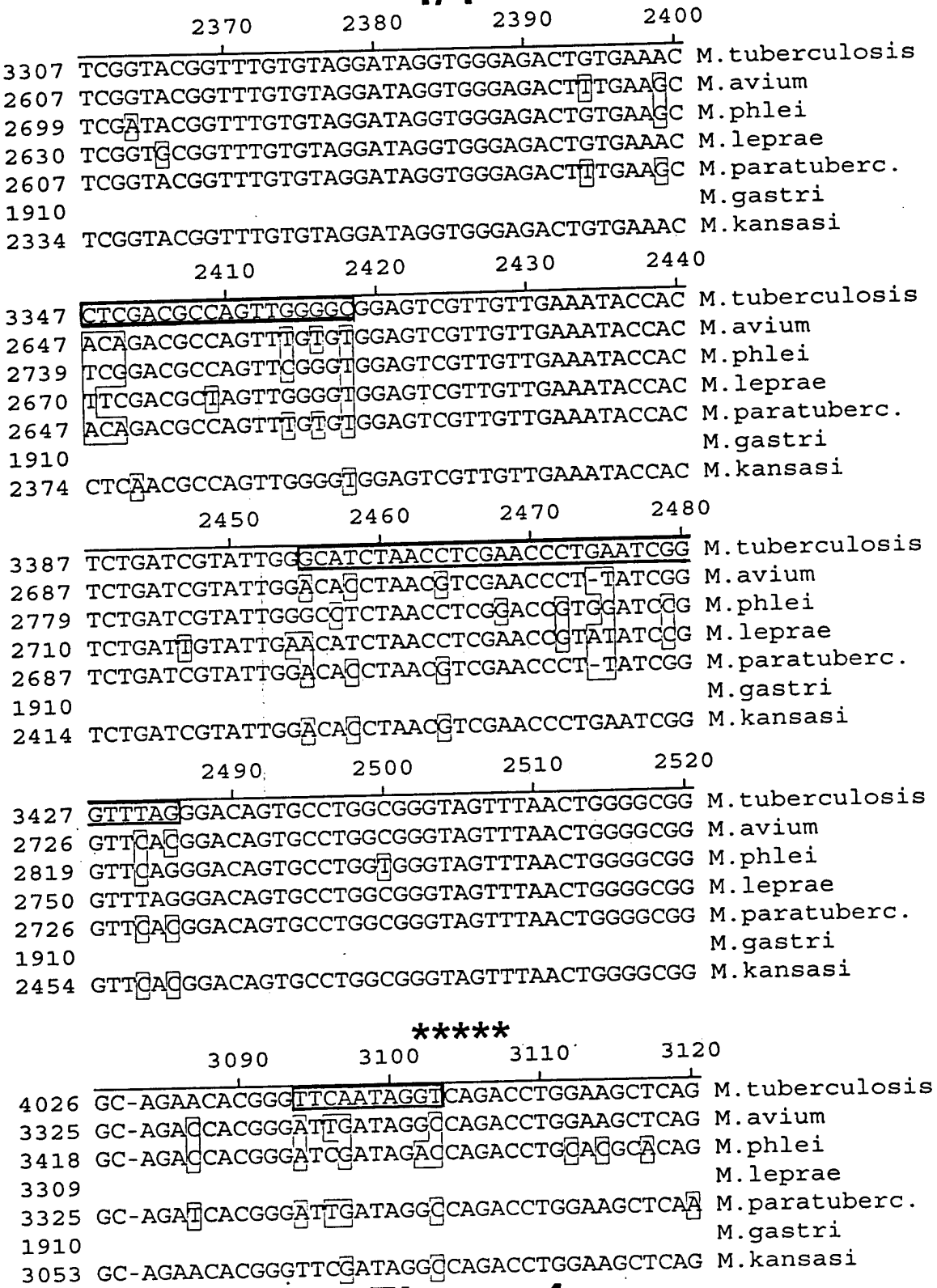

2370 2380 2390 2400
3307 TCGGTACGGTTTGTGTAGGATAGGTGGGAGACTGTGAAAC M.tuberculosis
2607 TCGGTACGGTTTGTGTAGGATAGGTGGGAGACTTTGAAGC M.avium
2699 TCGATACGGTTTGTGTAGGATAGGTGGGAGACTGTGAAGC M.phlei
2630 TCGGTGCGGTTTGTGTAGGATAGGTGGGAGACTGTGAAAC M.leprae
2607 TCGGTACGGTTTGTGTAGGATAGGTGGGAGACTTTGAAGC M.paratuberc.
1910 M.gastri
2334 TCGGTACGGTTTGTGTAGGATAGGTGGGAGACTGTGAAAC M.kansasi
2410 2420 2430 2440
3347 CTCGACGCCAGTTGGGGCGGAGTCGTTGTTGAAATACCAC M.tuberculosis
2647 ACAGACGCCAGTTTGTGTGGAGTCGTTGTTGAAATACCAC M.avium
2739 TCGGACGCCAGTTCGGGTGGAGTCGTTGTTGAAATACCAC M.phlei
2670 TTCGACGCTAGTTGGGGTGGAGTCGTTGTTGAAATACCAC M.leprae
2647 ACAGACGCCAGTTTGTGTGGAGTCGTTGTTGAAATACCAC M.paratuberc.
1910 M.gastri
2374 CTCAACGCCAGTTGGGGTGGAGTCGTTGTTGAAATACCAC M.kansasi
2450 2460 2470 2480
3387 TCTGATCGTATTGGGCATCTAACCTCGAACCCTGAATCGG M.tuberculosis
2687 TCTGATCGTATTGGACACGCTAACGTCGAACCCT-TATCGG M.avium
2779 TCTGATCGTATTGGGCGTCTAACCTCGGACCGTGGATCCG M.phlei
2710 TCTGATTGTATTGAACATCTAACCTCGAACCGTATATCCG M.leprae
2687 TCTGATCGTATTGGACACGCTAACGTCGAACCCT-TATCGG M.paratuberc.
1910 M.gastri
2414 TCTGATCGTATTGGACACGCTAACGTCGAACCCTGAATCGG M.kansasi
2490 2500 2510 2520
3427 GTTTAGGGACAGTGCCTGGCGGGTAGTTTAACTGGGGCGG M.tuberculosis
2726 GTTCACGGACAGTGCCTGGCGGGTAGTTTAACTGGGGCGG M.avium
2819 GTTCAGGGACAGTGCCTGGTGGGTAGTTTAACTGGGGCGG M.phlei
2750 GTTTAGGGACAGTGCCTGGCGGGTAGTTTAACTGGGGCGG M.leprae
2726 GTTCACGGACAGTGCCTGGCGGGTAGTTTAACTGGGGCGG M.paratuberc.
1910 M.gastri
2454 GTTCACGGACAGTGCCTGGCGGGTAGTTTAACTGGGGCGG M.kansasi
*****
3090 3100 3110 3120
4026 GC-AGAACACGGGTTCAATAGGTCAGACCTGGAAGCTCAG M.tuberculosis
3325 GC-AGACCACGGGATTGATAGGCCAGACCTGGAAGCTCAG M.avium
3418 GC-AGACCACGGGATCGATAGACCAGACCTGCACGCACAG M.phlei
3309 M.leprae
3325 GC-AGATCACGGGATTGATAGGCCAGACCTGGAAGCTCAA M.paratuberc.
1910 M.gastri
3053 GC-AGAACACGGGTTCGATAGGCCAGACCTGGAAGCTCAG M.kansasi

Figure 4